DEVIEW
2017

책 읽어주는 딥러닝:

배우 유인나가 해리포터를 읽어준다면

김태훈 / carpedm20

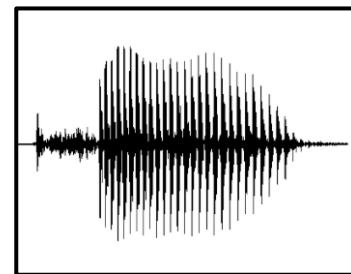
DEVSISTERS

김태훈 / carpedm20 / DEVSISTERS

오늘의 주제

음성 합성

텍스트 → 음성 합성 → 음성

"이거 실화냐?" → 음성 합성 →

무엇을 할 수 있나요?

1. 음성 안내 시스템 지하철, 박물관
2. 대화 인공지능 Siri, 스피커
3. 오디오북 성우가 필요한 대부분의 일

많은 **기업**들이 기술을 보유

음성합성
오디오북 듣기

세상 하나뿐인 오디오북, 유인나가 먼저 읽고 나머지는 기술이 완성합니다

YG와 네이버랩스가 함께하는 〈오디오북 기부 캠페인〉

[" 모두가 책과 더 가까워 지기를 바라며 특별한 오디오북을 준비했어요.
앞부분은 제가 정성껏 녹음했고, 나머지는 네이버 음성합성기술이
제 목소리를 재현해서 완성했답니다. 세상 하나뿐인 오디오북, 어떻게 들리시나요?"]

유인나 데모를 듣고나서..

"배우 유인나가 해리포터를 읽어준다면?"

나도 할 수 있을까?

API

kakao NAVER

유인나는 없음 ㅠㅠ

Clova Speech Synthesis (CSS)

텍스트를 성우의 음성으로 자연스럽게 읽어주는 음성 합성 API

아름다운 성우의 목소리로 글을 읽어주는 애플리케이션을 만들 때 유용한 API입니다. 주어진 텍스트를 자연스럽게 말하듯 재생해주어 음성 안내 시스템, 뉴스/책 읽기 서비스 등에 활용할 수 있습니다.

자연스러운 합성음

네이버에서 수년간 연구해온 결과로 국내에서 가장 자연스러운 합성음을 제공합니다. Clova에 적용된 음성 합성 기술과 동일한 수준의 API를 사용할 수 있으며, 한국어 외에도 영어, 일어, 중국어를 지원합니다.

https://developers.naver.com/products/clova/tts/

Clova Speech Synthesis (CSS)

텍스트를 성우의 음성으로 자연스럽게 읽어주는 음성 합성 API

아름다운 성우의 목소리로 글을 읽어주는 애플리케이션을 만들 때 유용한 API입니다. 주어진 텍스트를 자연스럽게 말하듯 재생해주어 음성 안내 시스템, 뉴스/책 읽기 서비스 등에 활용할 수 있습니다.

자연스러운 합성음

네이버에서 수년간 연구해온 결과로 국내에서 가장 자연스러운 합성음을 제공합니다. Clova에 적용된 음성 합성 기술과 동일한 수준의 API를 사용할 수 있으며, 한국어 외에도 영어, 일어, 중국어를 지원합니다.

https://developers.naver.com/products/clova/tts/

Clova Speech Synthesis (CSS)

텍스트를 성우의 음성으로 자연스럽게 읽어주는 음성 합성 API

아름다운 성우의 목소리로 글을 읽어주는 애플리케이션을 만들 때 유용한 API입니다. 주어진 텍스트를 자연스럽게 말하듯 재생해주어 음성 안내 시스템, 뉴스/책 읽기 서비스 등에 활용할 수 있습니다.

자연스러운 합성음

네이버에서 수년간 연구해온 결과로 국내에서 가장 자연스러운 합성음을 제공합니다. Clova에 적용된 음성 합성 기술과 동일한 수준의 API를 사용할 수 있으며, 한국어 외에도 영어, 일어, 중국어를 지원합니다.

https://developers.naver.com/products/clova/tts/

음..

아쉽다

“자연스러운 목소리를 마음껏”

그래서 로 **직접** 만들어 봤습니다

한번 들어볼까요?

"존경하는 개발자 여러분."

“저는 데브시스터즈에서

머신러닝 엔지니어로 일하고 있는 김태훈입니다.”

"이렇게 제 발표를 들으러 와 주셔서
정말 감사합니다."

"아직은 턱없이 부족한

음성합성 모델이지만"

"오늘 데이터와 코드를 오픈소스로 공개해서"

"여러분들과 함께 더 좋은 모델로

만들어 나갔으면 좋겠습니다"

“코드는 발표가 끝나고 공개할 예정이며,”

"이제 다시 발표자가 진행하도록 하겠습니다."

어떠신가요?

오늘의 발표

1. 데이터
2. 모델
3. 결과

DEVSISTERS

1. 데이터

2. 모델

3. 결과

DEVSISTERS

텍스트 ↔ 음성

공개 데이터가 없다

직접 만들어야..

kakao

뉴스검색

뉴스홈 주요뉴스 이슈 랭킹뉴스 TV 뉴스통계

9.3°C 전주 다음앱

일간 | 주간 | 월간 | 기자댓글 | 앵커브리핑

손석희 앵커의 댓글을 이제 음성으로 들어보세요

손석희의 앵커브리핑

손석희앵커
2015.04.10 09:28

소설가 마크 트웨인이 말했습니다.
"인생에 가장 중요한 이틀이있는데, 하나는 세상에 태어난 날이고 다른 하나는 왜 이 세상에 왔는가를 깨닫는 날이다."
그런데 그 첫번째 날은 누구나 다 알지만, 두번째 날은 참 어려운 것 같습니다. 누구나 그 두번째 날을 만나기 위해 애쓰는게 삶인지도 모르겠습니다.
뉴스룸도 그런 면에서 똑같습니다. 저희들도 그 두번째의 날을 만나고 기억하기 위해 매일 매일 최선을 다하겠습니다.

손석희앵커
2014.08.27 16:56

세상은 세월호 특별법을 가지고 혼란스럽습니다. 세월호 참사의 본질은 무엇인가를 다시 생각하게 하는 요즈음입니다. 진상규명이든 보상이든 특별법을 가지고 이렇게 보혁갈등을 일으키는 사회가 정상적인가를 생각해봐야 하지 않을까 합니다. 세월호 참사에 대해 저희는 아직도 왜?라는 질문을 버리지 않고 있습니다. 즉, 참사의 원인에 대해 규명하지 않고서는 한 걸음도 더 나아갈 수 없다고 믿기 때문입니다. 저희는 이런 작업이 음모론에 치우치는 것도 원치 않습니다. 단지 취재한 것을 사실 그대로 전해드리는 것이 저희들의 원칙이고 그것이 저널리즘의 역할이라고 믿습니다. 오늘도 왜?라는 질문에 대해 팽목항에서 김관 기자가 전해드릴 내용들이 있습니다. 이따가 뵙겠습니다.

JMnet 회원가입 로그인

드라마 예능 교양 시사 | 편성표 온에어

뉴스홈 다시보기 오픈 저널리즘 JTBC 뉴스룸

포토뉴스 영상구성 랭킹 APP

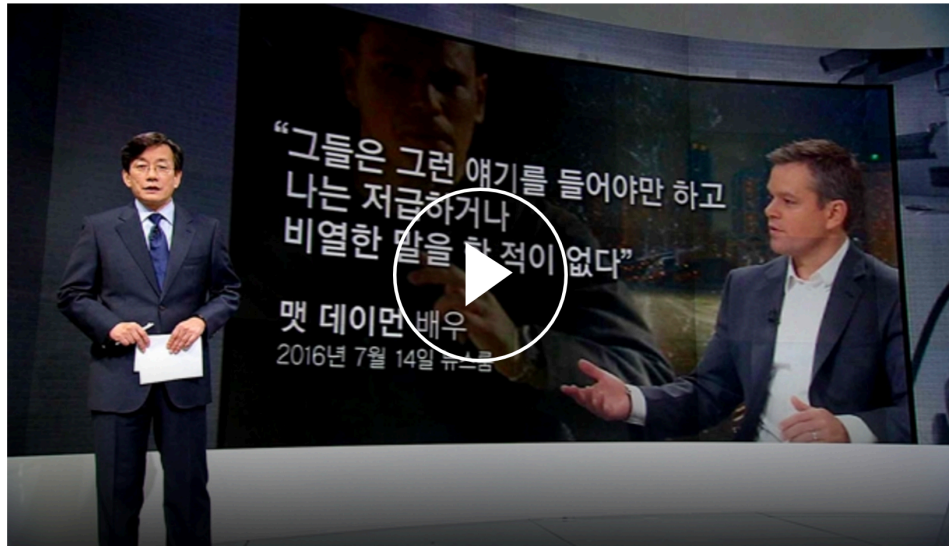
JTBC 뉴스룸 | 아침& | 뉴스 현장 | 정치부 회의 | 보도특집

여러분의 제보를 기다립니다!

JTBC, 3년 연속 시청자가 뽑은 최고의 방송사 선정!

〔앵커브리핑〕 '나는 비열하거나 저급하지 않았다'

[JTBC] 입력 2017-09-12 21:39

안내▸ JTBC 뉴스는 여러분의 생생한 제보를 기다리고 있습니다.

뉴스룸의 앵커브리핑을 시작하겠습니다.

이명박 정부 당시 배우 김여진 씨는 정치사회적으로 할 말은 해서 이른바 '개념 배우'라는 별명을 얻고 있었습니다.

연예인이 정치적 발언을 하면 왜 개념이 있다는 칭찬을 들어야 하는지…그것도 어찌 보면 한국적 상황인지는 모르겠습니다.

아무튼 배우 김여진 씨를 전원책 변호사의 맞상대로 해서 토론 코너에 출연시키려던 당시 라디오 프로그램의 시도는 무산됐습니다. 그 라디오 프로그램은 제가 진행하던 프로그램이었습니다.

이 시도는 급기야 소셜테이너 출연 금지 규정이라는 것까지 사내에 생겨나게 했지요. 정치적 입장을 가진 연예인은 방송에 출연시킬 수 없다는 것이었습니다.

다시 말하면 연예인은 정치적 입장을 공개적으로 또한 지속적으로 표현할 수 없다는 것에 다름이 아니었습니다.

비슷한 시기에 김미화 씨는 코미디언으로서 시사 프로그램을 진행하는 것에 대해서 여기저기서 격려도 받았지만, 동시에 비웃음도 들어야 했습니다.

이것도 사실은 시사 프로그램에 코미디언이 진행을 맡아서는 안 된다는 비뚤어진 편견의 소산이었습니다.

http://news.jtbc.joins.com/html/033/NB11521033.html

데이터 프로세싱

1. 음성 추출

+

2. 문장 별 자르기

+

3. 텍스트 ↔ 음성 맞추기

노가다 노가다 노가다 노가다 노가다
노가다 노가다 노가다 노가다 노가다
노가다 노가다 노가다 노가다 노가다
노가다 노가다 노가다 노가다 노가다
노가다 노가다 노가다 노가다 노가다
노가다 노가다 노가다 노가다 노가다
노가다 노가다 노가다 노가다 노가다

Google Speech API, Text similarity

직접 하진 않았고 자동화

뉴스 + YouTube + 오디오북

체감상 90% 정확도

총 5명의 목소리

완벽하진 않지만 50+시간 데이터

손석희 :

손석희 : 15+시간 13,000 문장

손석희 : 15+시간 13,000 문장

박근혜 : 5+ 시간 6,000문장

손석희 : 15+시간 13,000 문장

박근혜 : 5+ 시간 6,000문장

문재인 : 2+ 시간 2,000 문장

1. 데이터

2. 모델

3. 결과

DEVSISTERS

딥러닝

총 2가지 모델

1. Tacotron
2. Deep Voice 2

1. Tacotron

7달 전

TACOTRON: TOWARDS END-TO-END SPEECH SYNTHESIS

Yuxuan Wang[*], RJ Skerry-Ryan[*], Daisy Stanton, Yonghui Wu, Ron J. Weiss[†], Navdeep Jaitly,

Zongheng Yang, Ying Xiao[*], Zhifeng Chen, Samy Bengio[†], Quoc Le, Yannis Agiomyrgiannakis,

Rob Clark, Rif A. Saurous[*]

Google, Inc.
{yxwang,rjryan,rif}@google.com

[*]These authors really like tacos.
[†]These authors would prefer sushi.

핵심은

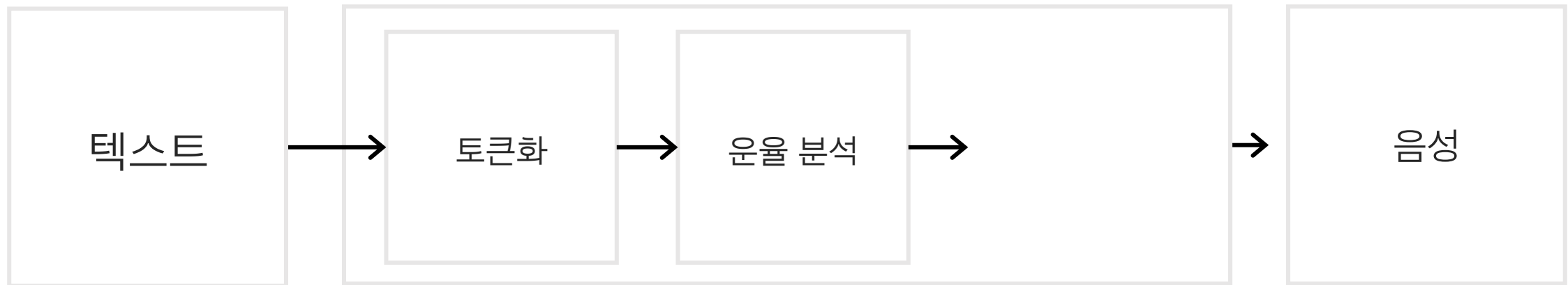
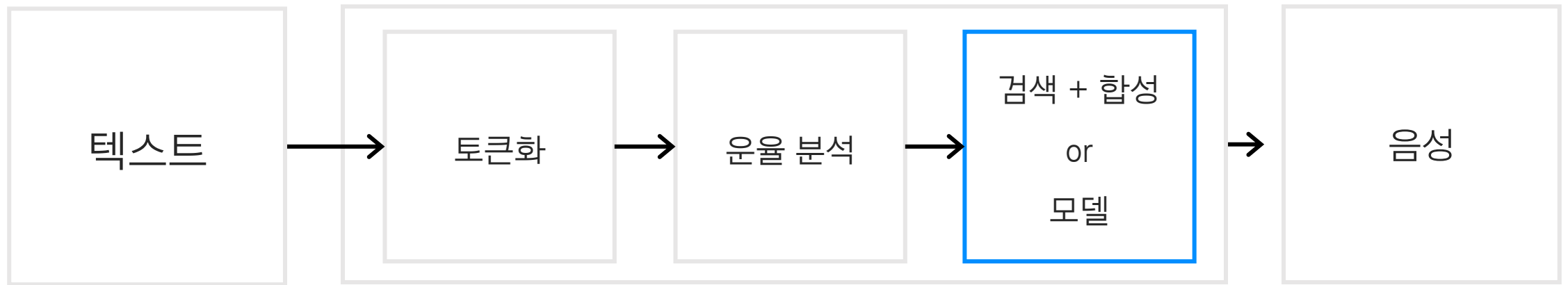
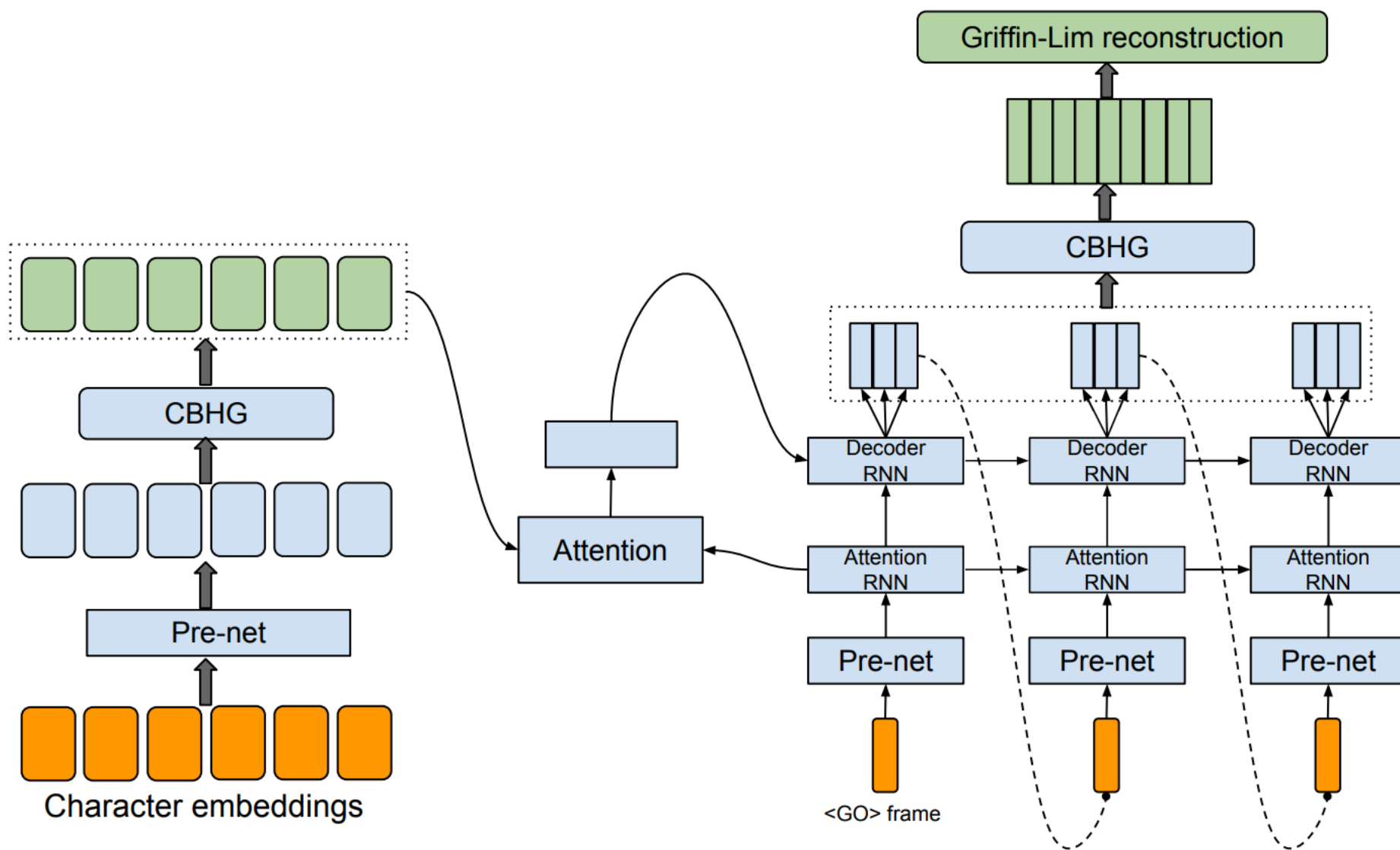
텍스트
음성

텍스트
토큰화
운율 분석
음성

텍스트
토큰화
운율 분석
검색 + 합성
or
모델
음성

텍스트
모델
음성

Tacotron

Wang, Yuxuan, et al. "Tacotron: A Fully End-to-End Text-To-Speech Synthesis Model." arXiv preprint arXiv:1703.10135 (2017).

크게 4가지 모듈로 구성

1. Encoder

2. Decoder

Decoder RNN
Decoder RNN
Decoder RNN
Attention RNN
Attention RNN
Attention RNN
Pre-net
Pre-net
Pre-net
<GO> frame

3. Attention

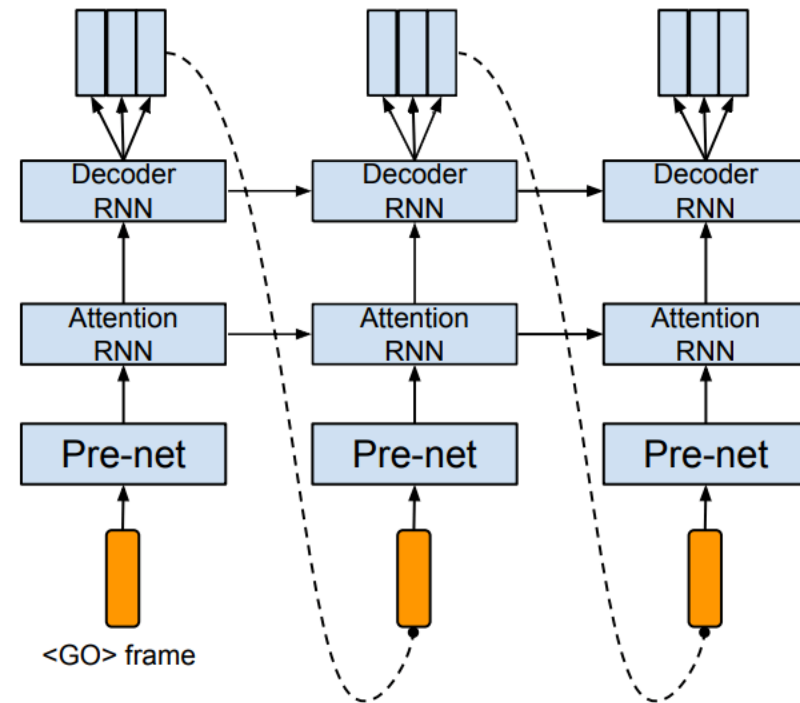
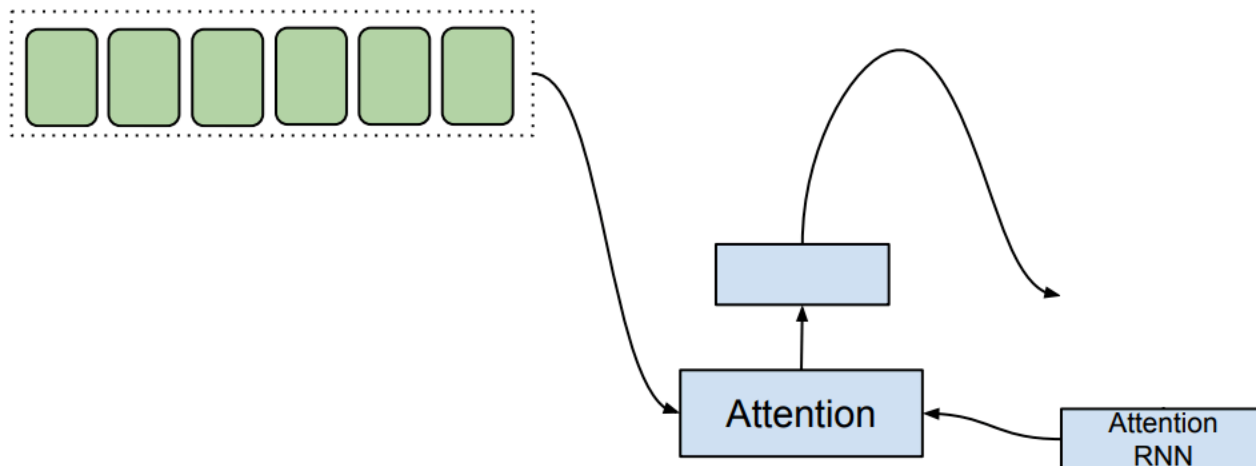
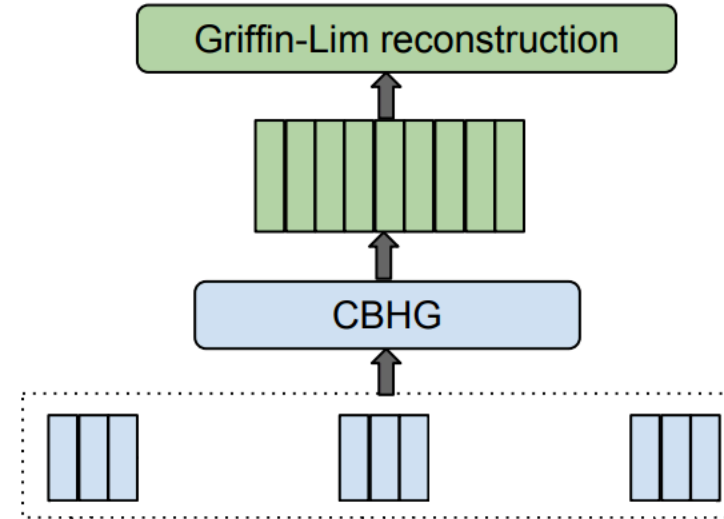
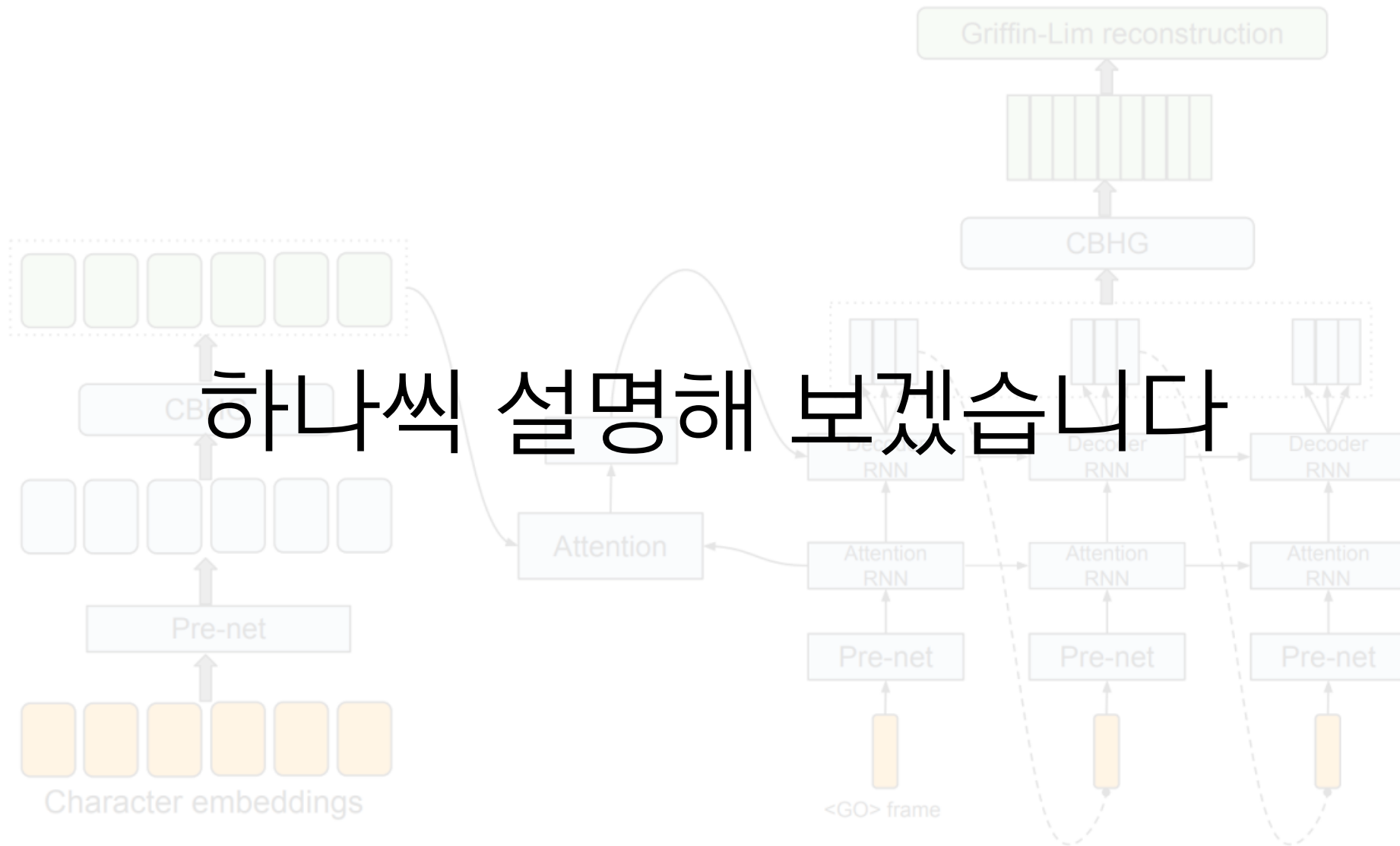
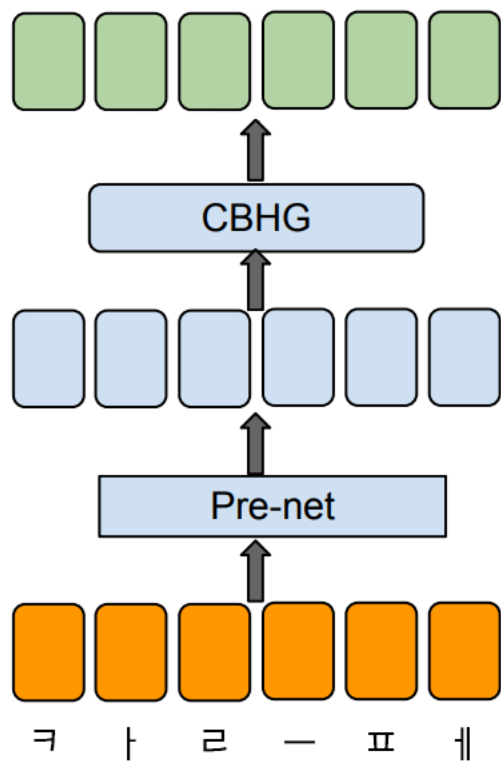
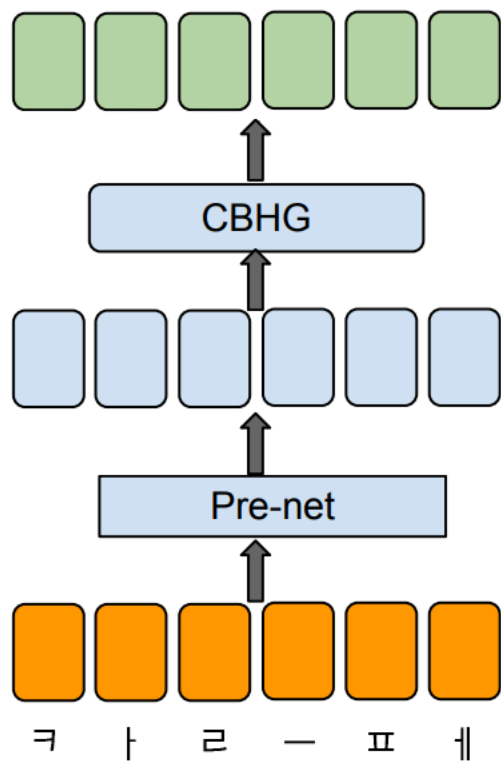
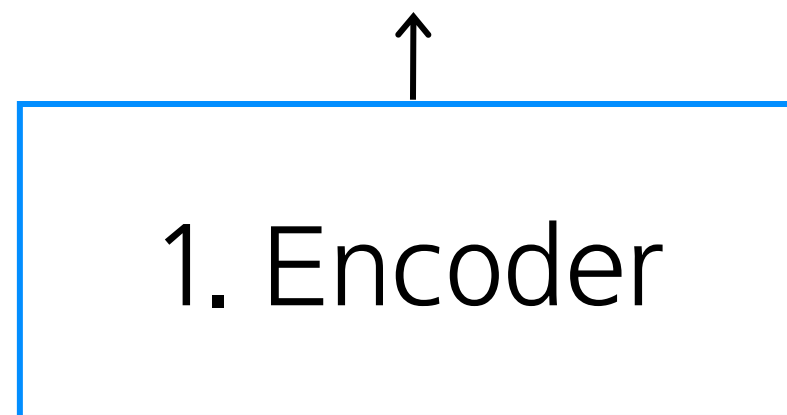
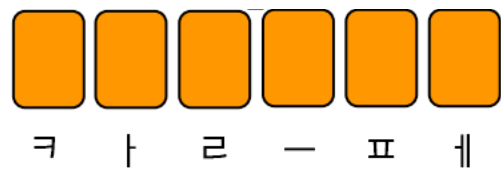
4. Vocoder

하나씩 설명해 보겠습니다

1. Encoder

CBHG
Pre-net
ㅋ ㅏ ㄹ ㅡ ㅍ ㅔ

텍스트 정보를 잘 나타내는 숫자
1. Encoder
텍스트

ㅋ ㅏ ㄹ ㅡ ㅍ ㅔ

Character embeddings

Character Embedding이란?

ㄱ, ㄴ, ㅏ, ㅑ

글자 →

0.1, 0.3

글자 →숫자

딥러닝 모델이 **계산**을하기 위해

글자 → 숫자

ㄱ
ㄴ
ㅏ
ㅑ
0.7
0.5
0.9
0.5
1.0
0.9
...
0.3
0.0
0.7
0.5
0.9
0.9
...
0.1
0.0
0.3
0.0
0.0
0.2
...
0.9
0.5
0.1
0.1
0.0
0.2
...

Character Embedding의 장점

띵작

띵작

명작

띵작

명작
0.0 0.0 0.1
0.9 0.0 0.1
0.8 0.0 0.2
0.5 1.0 0.9
0.7 0.5 0.9
0.3 0.0 0.2
ㅁ ㅕ ㅇ ㅈ ㅏ ㄱ
띵동
0.1 0.0 0.5
0.1 0.0 0.5
0.5 0.1 0.2
0.1 0.0 0.5
0.0 0.9 0.5
0.8 0.0 0.2
ㄷ ㄷ ㅣ ㄷ ㅗ ㅇ

띵작

0.1 0.0 0.5
0.1 0.0 0.5
0.5 0.1 0.2
ㄷ ㄷ ㅣ

명작

0.0 0.0 0.1
0.9 0.0 0.1
0.8 0.0 0.2
0.5 1.0 0.9
0.7 0.5 0.9
0.3 0.0 0.2
ㅁ ㅕ ㅇ ㅈ ㅏ ㄱ

띵동

0.1 0.0 0.5
0.1 0.0 0.5
0.5 0.1 0.2
0.1 0.0 0.5
0.0 0.9 0.5
0.8 0.0 0.2
ㄷ ㄷ ㅣ ㄷ ㅗ ㅇ

ㄱ ㄲ ㄴ ㄷ ㄸ ㄹ ㅁ ㅂ ㅃ ㅅ ㅆ ㅇ ㅈ ㅉ ㅊ ㅋ ㅌ ㅍ ㅎ

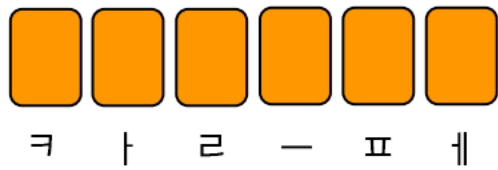
ㅏ ㅐ ㅑ ㅒ ㅓ ㅔ ㅕ ㅖ ㅗ ㅘ ㅙ ㅚ ㅛ ㅜ ㅝ ㅞ ㅟ ㅠ ㅡ ㅢ ㅣ

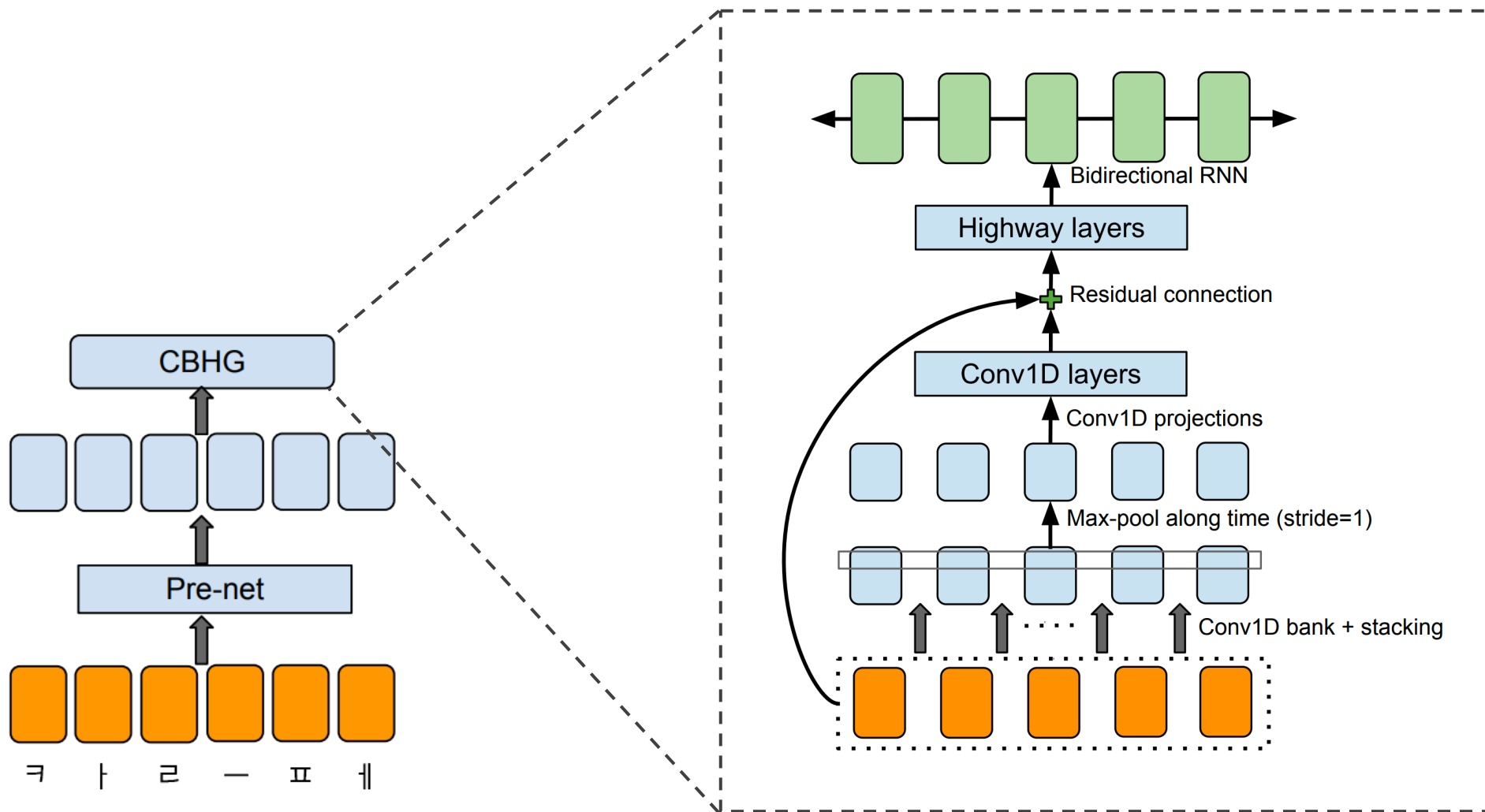
ㄱ ㄲ ㄳ ㄴ ㄵ ㄶ ㄷ ㄹ ㄺ ㄻ ㄼ ㄽ ㄾ ㄿ ㅀ ㅁ ㅂ ㅄ ㅅ ㅆ ㅇ ㅈ ㅊ ㅋ ㅌ ㅍ ㅎ

! ' ( ) , - . : ; ? _ ~

총 80개의 Embedding

Character embeddings

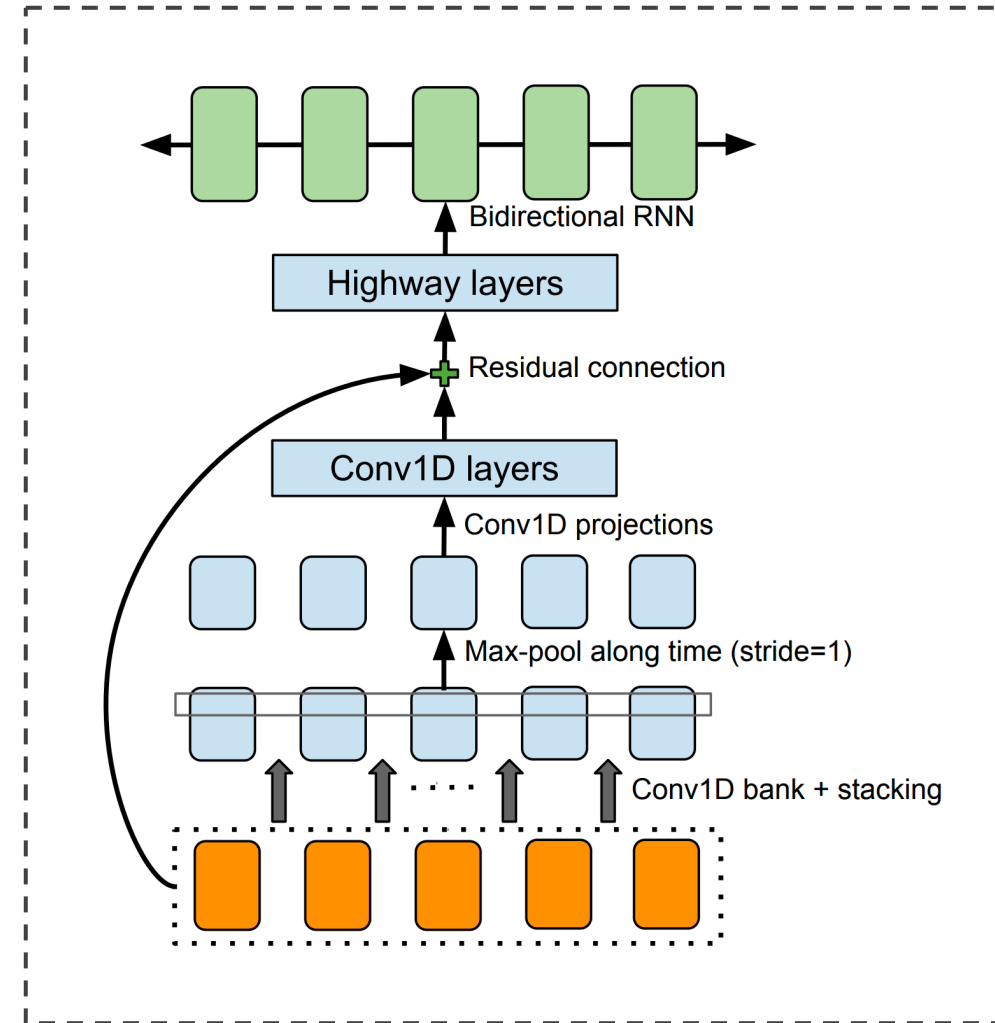
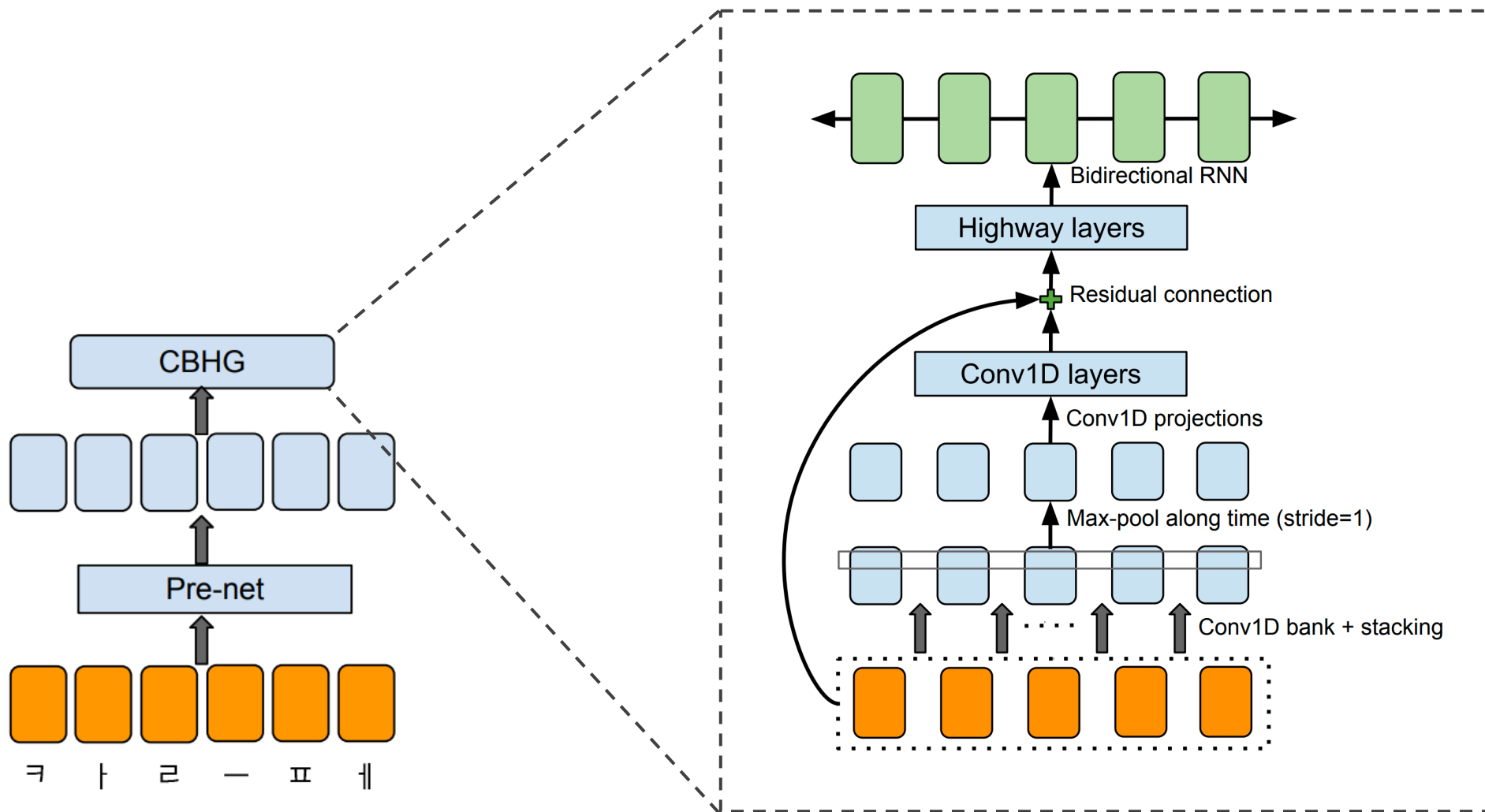
ㅋ ㅏ ㄹ ㅡ ㅍ ㅔ

FC – ReLU – Dropout – FC – ReLU – Dropout
Pre-net
ㅋ ㅏ ㄹ ㅡ ㅍ ㅔ

CBHG
Pre-net
ㅋ ㅏ ㄹ ㅡ ㅍ ㅔ
Bidirectional RNN
Highway layers
Residual connection
Conv1D layers
Conv1D projections
Max-pool along time (stride=1)
Conv1D bank + stacking

CBHG =

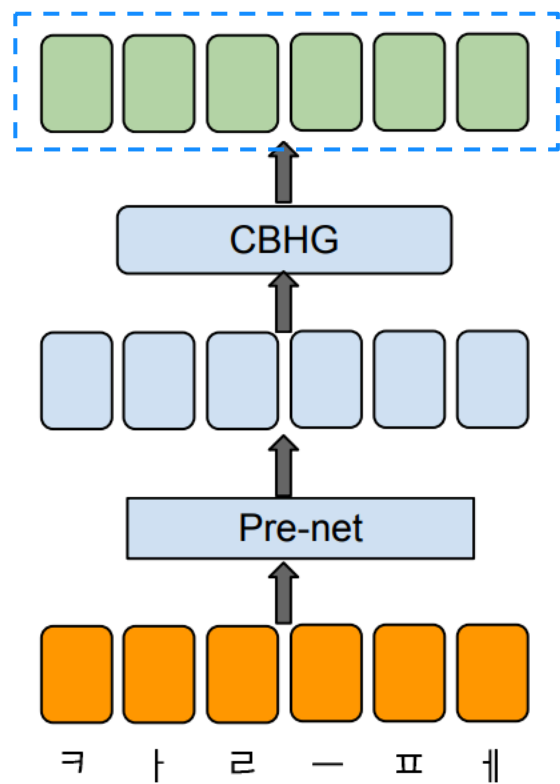
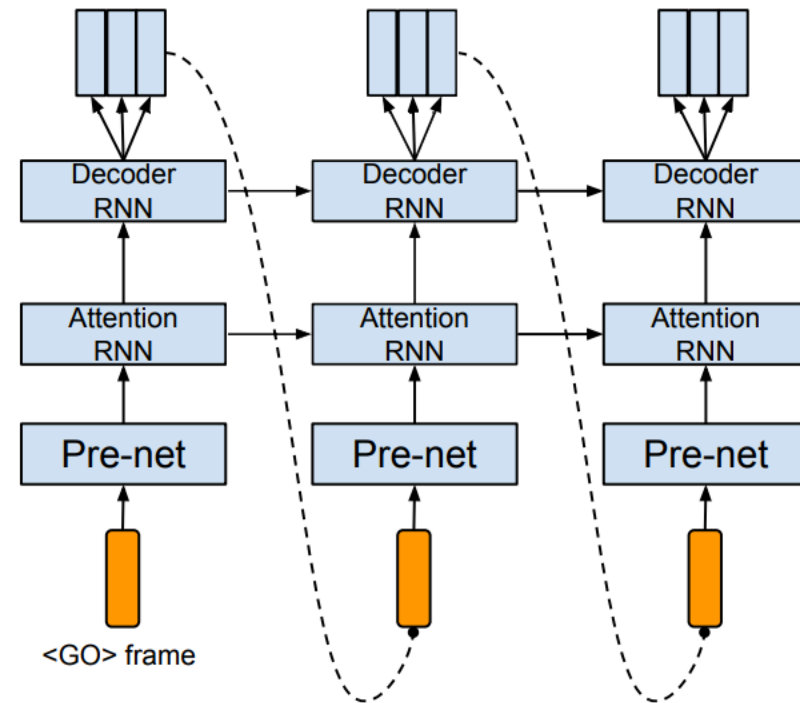
1-D Convolution Bank +

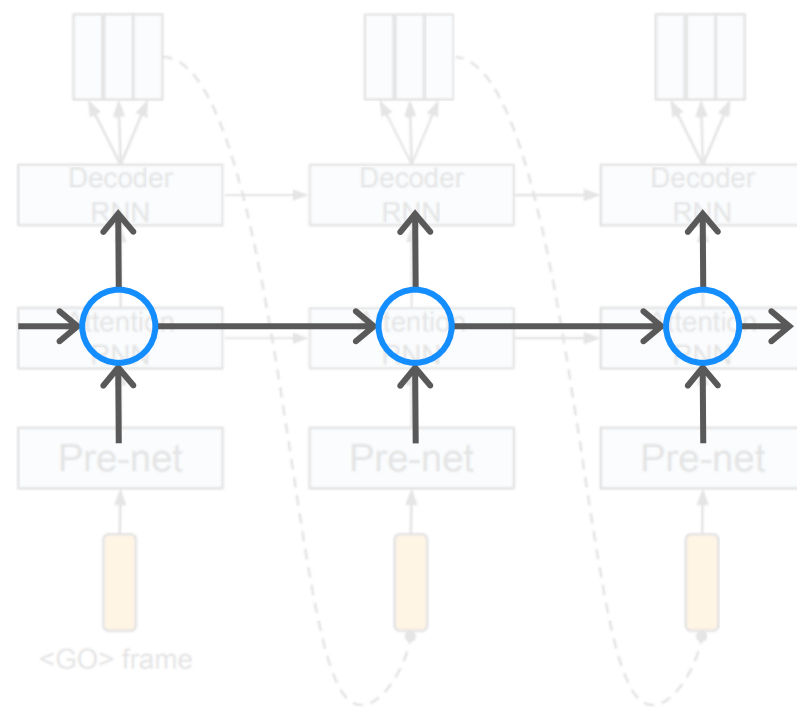
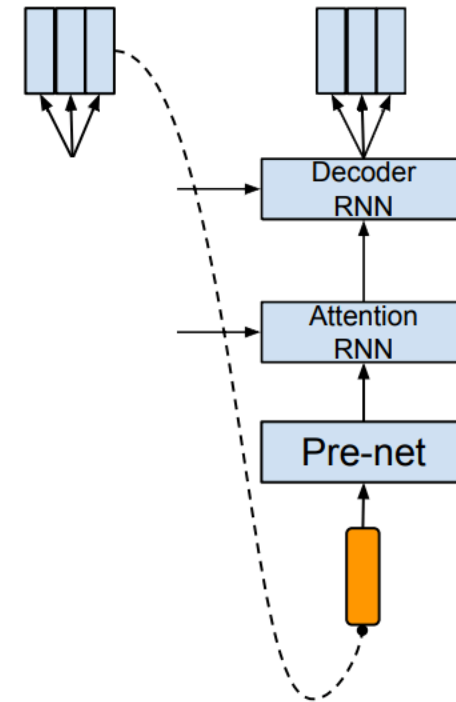
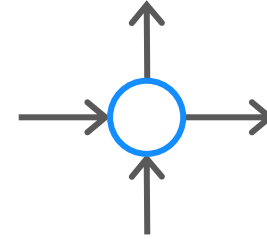
Highway network +

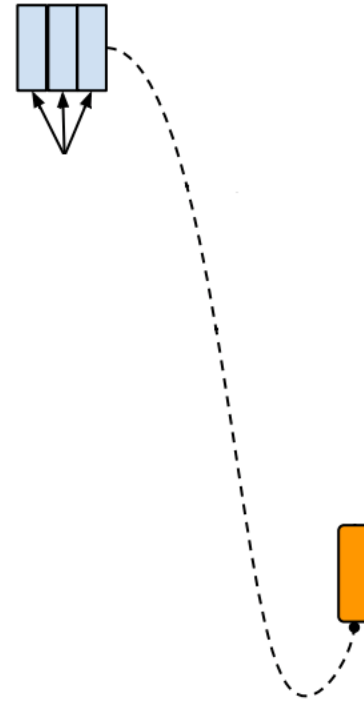
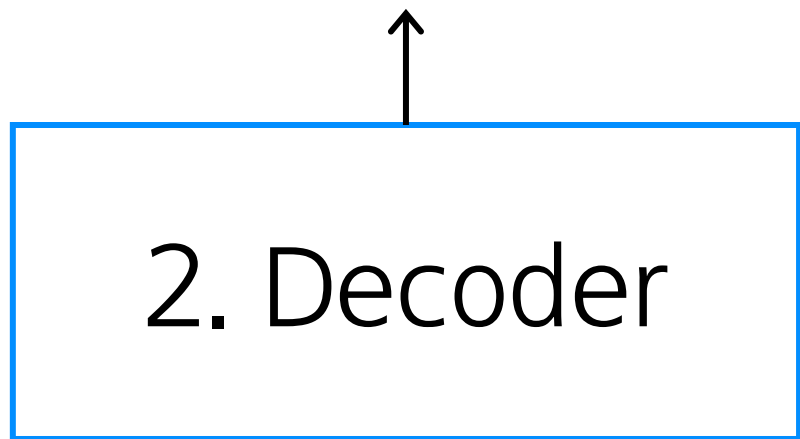
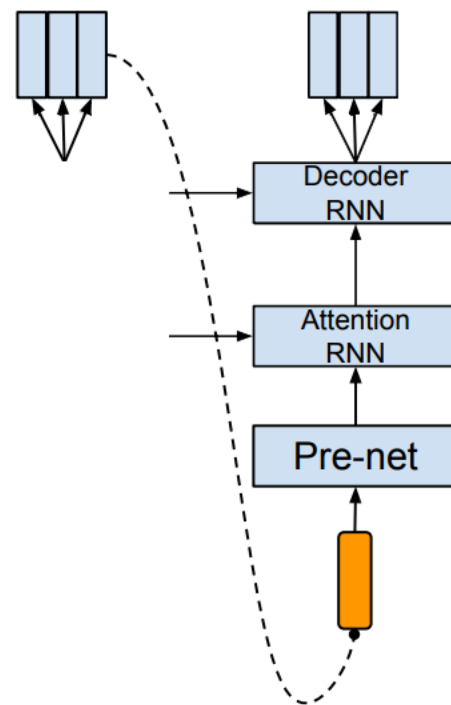
Bidirectional GRU

CBHG
Pre-net
ㅋ
ㅏ
ㄹ
ㅡ
ㅍ
ㅔ
Bidirectional RNN
Highway layers
Residual connection
Conv1D layers
Conv1D projections
Max-pool along time (stride=1)
Conv1D bank + stacking

텍스트 임베딩 : 텍스트를 잘 나타내는 숫자

2. Decoder

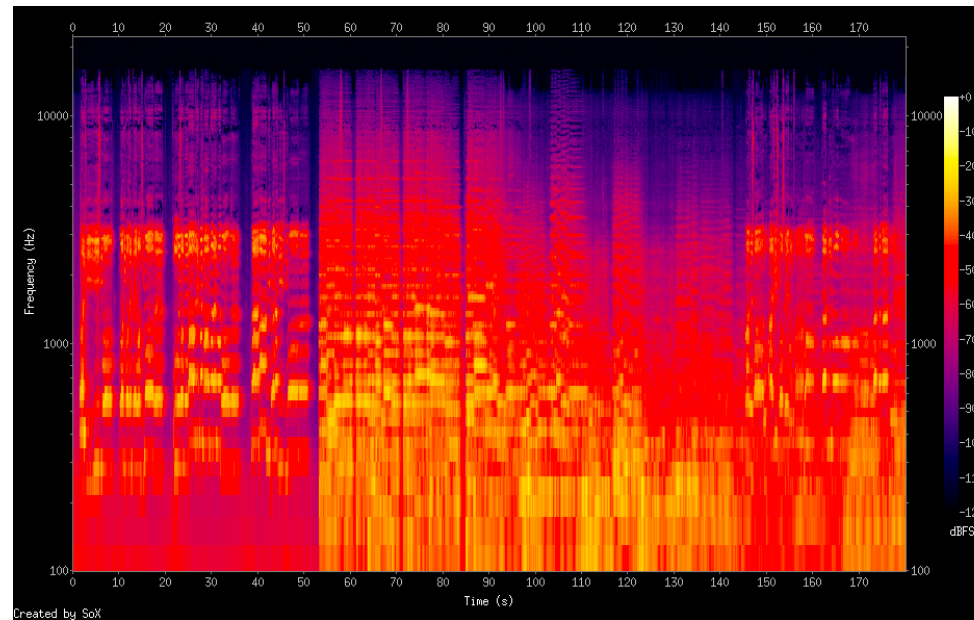
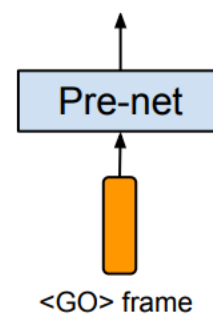
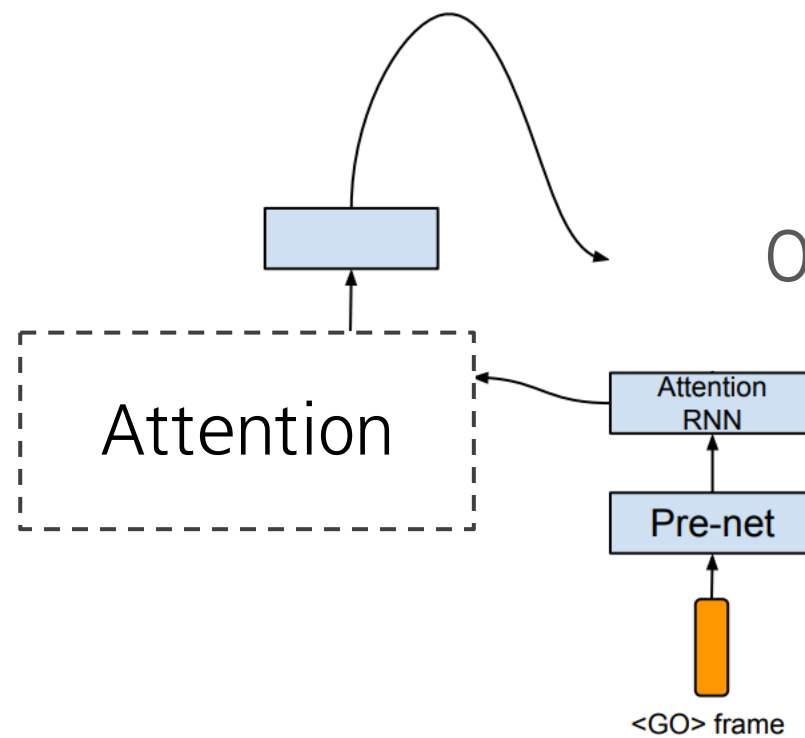
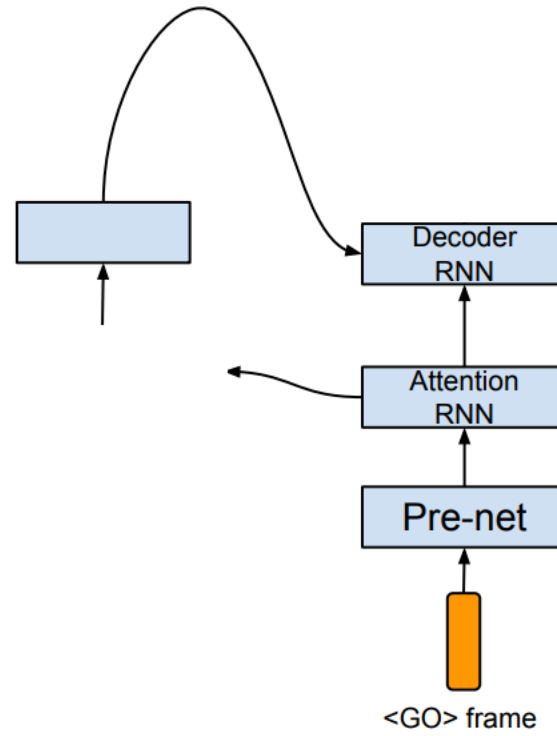
Decoder RNN
Decoder RNN
Decoder RNN
Attention RNN
Attention RNN
Attention RNN
Pre-net
Pre-net
Pre-net
<GO> frame

크게 보면 그냥 RNN

RNN Cell 하나

이전에 만든 스펙트로그램

다음 스팩트로그램
2. Decoder
이전에 만든 스팩트로그램

Decoder RNN
Attention RNN
Pre-net

스펙트로그램으로 음성을 만들 수 있다

문장에 상관없이 항상 같은 값

Pre-net
<GO> frame

이후 설명할 Attention
Attention
Attention
RNN
Pre-net
<GO> frame

그냥 RNN

예측해야하는 횟수를 1/N로

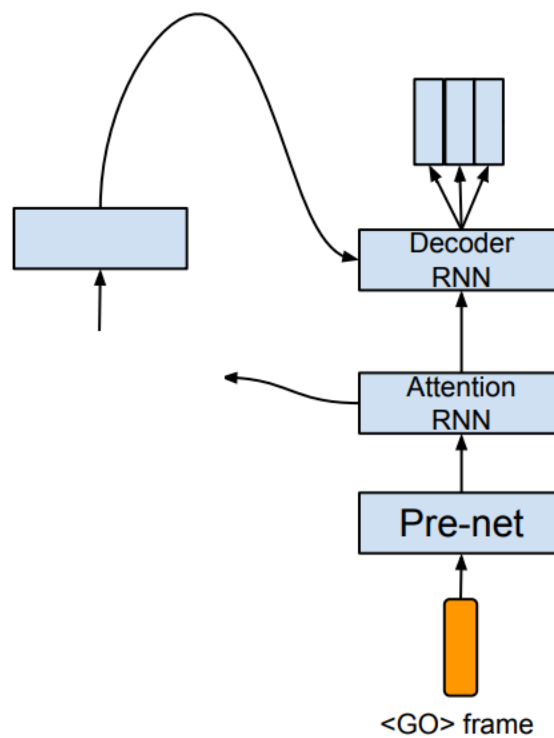
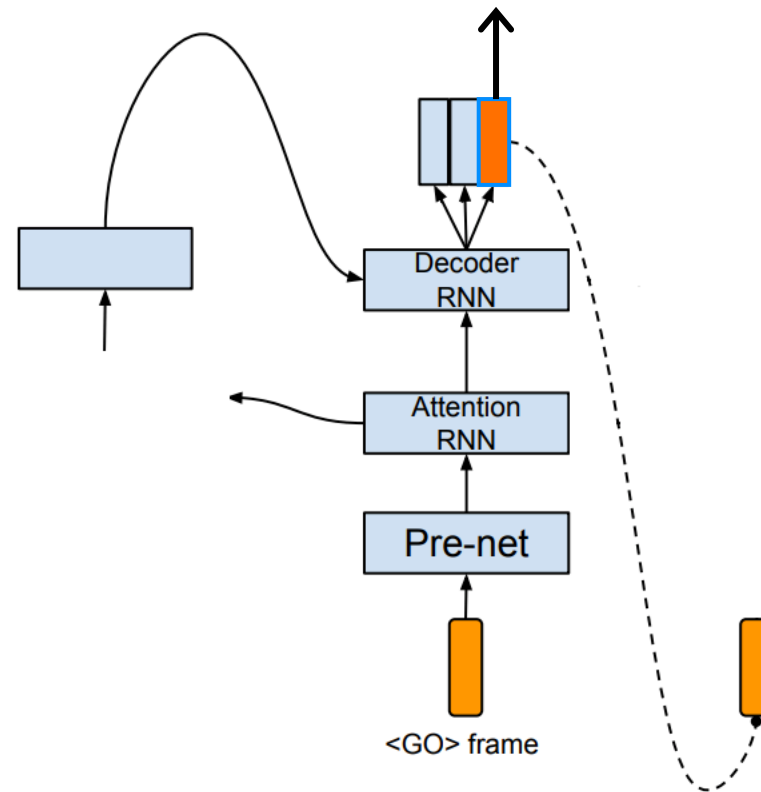
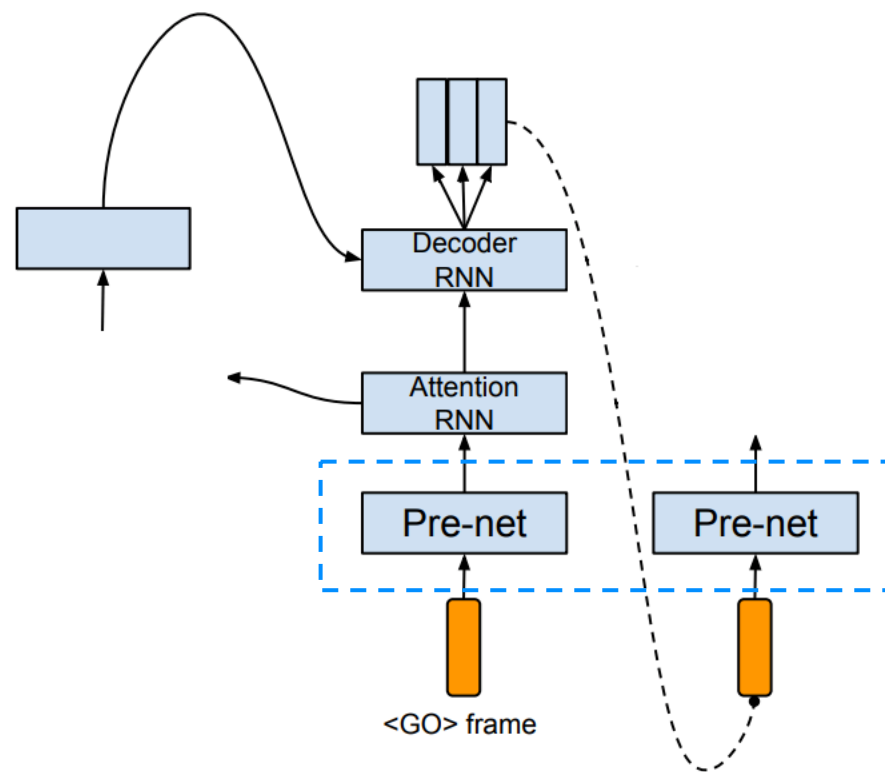
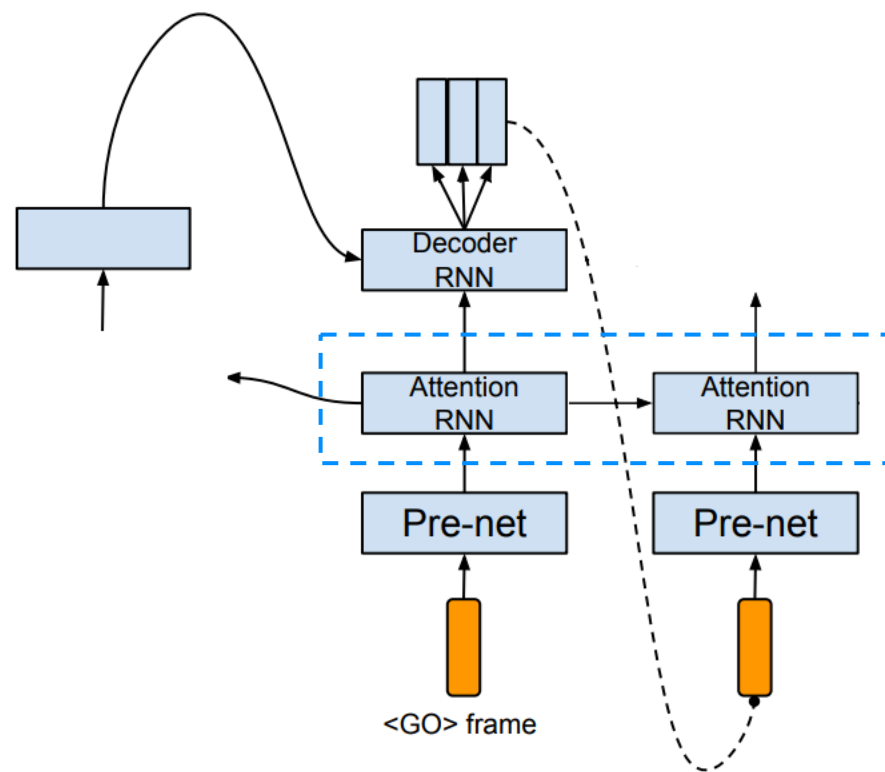
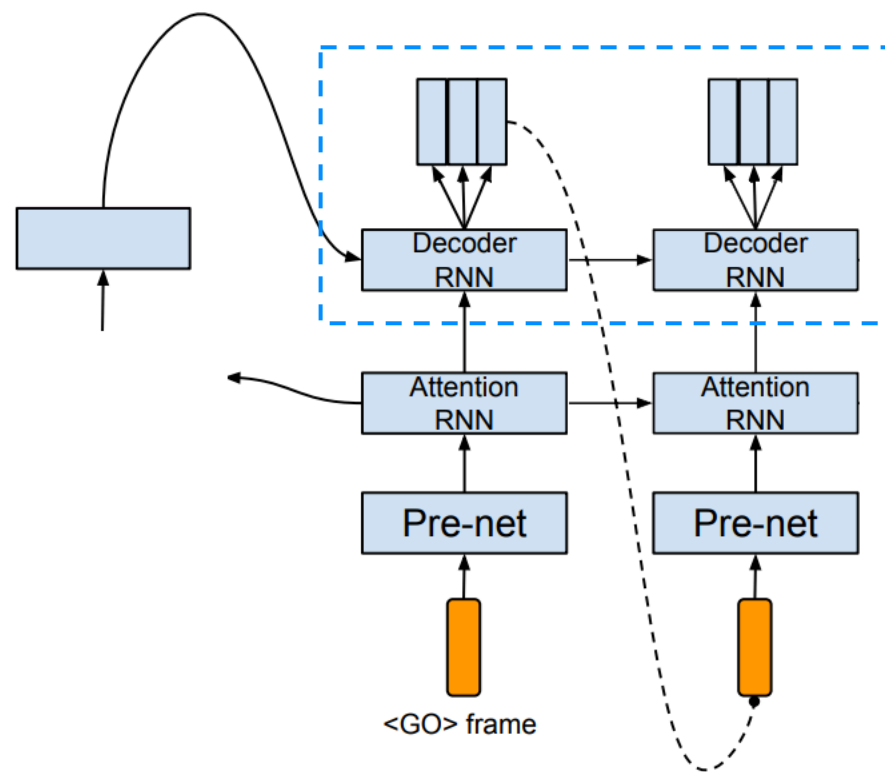
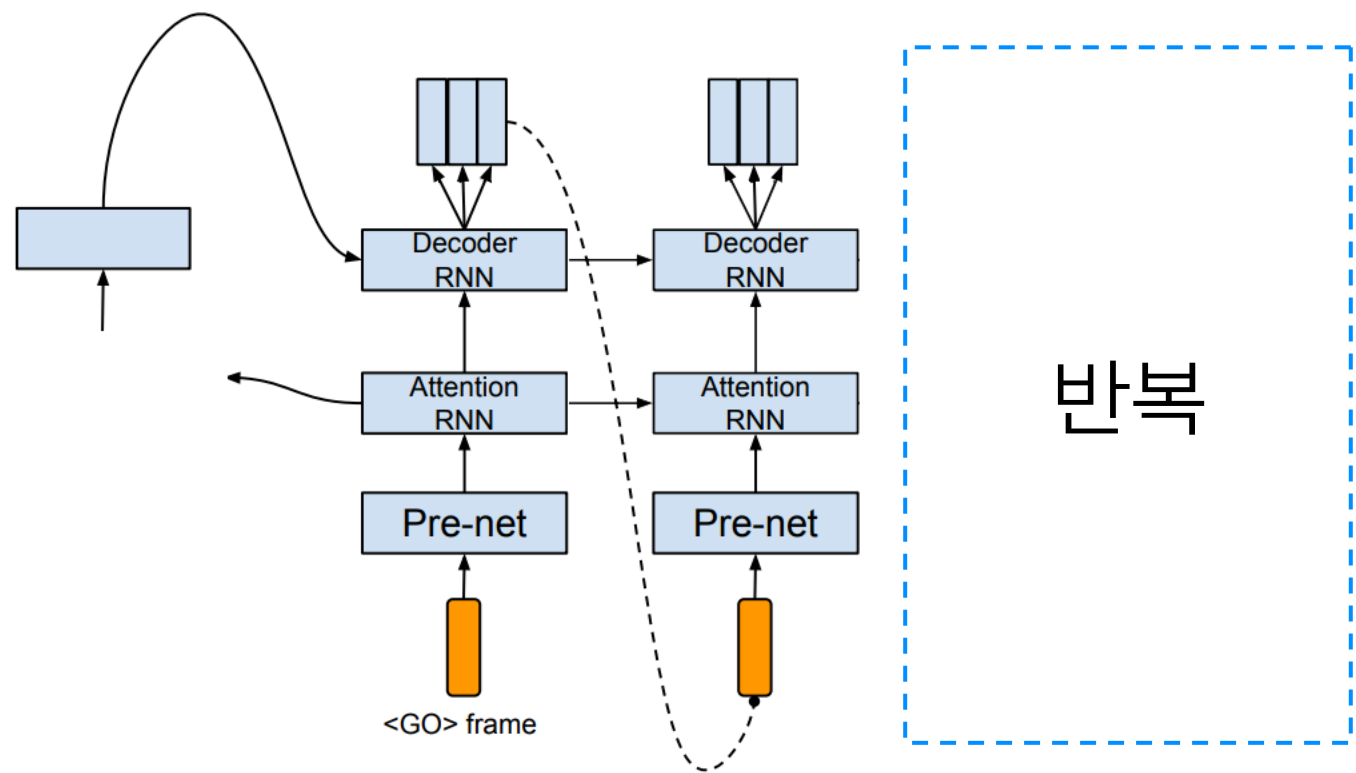
n개의 스펙트로그램 예측

마지막 스펙트로그램이 다음 입력값

Decoder RNN
Attention RNN
Pre-net
Pre-net
<GO> frame
똑같고

Decoder RNN
Attention RNN
Attention RNN
Pre-net
Pre-net
<GO> frame
똑같고

Decoder RNN
Decoder RNN
Attention RNN
Attention RNN
Pre-net
Pre-net
<GO> frame
똑같고

Decoder
RNN
Decoder
RNN
Attention
RNN
Attention
RNN
Pre-net
Pre-net
<GO> frame
반복

음성이 되기 직전의 숫자들

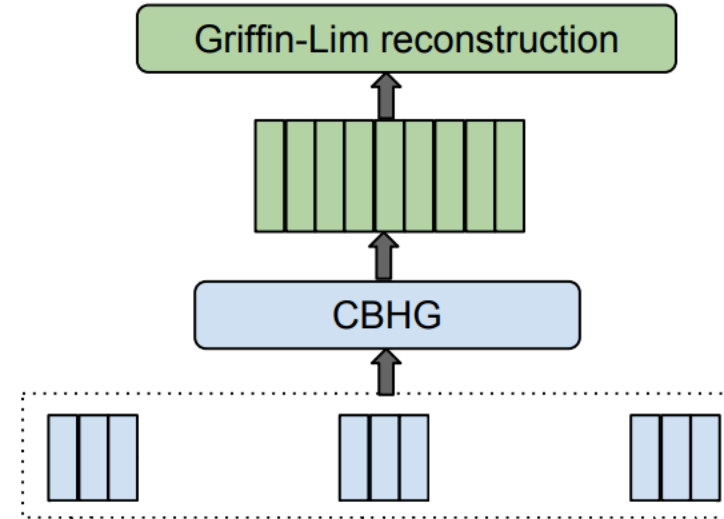
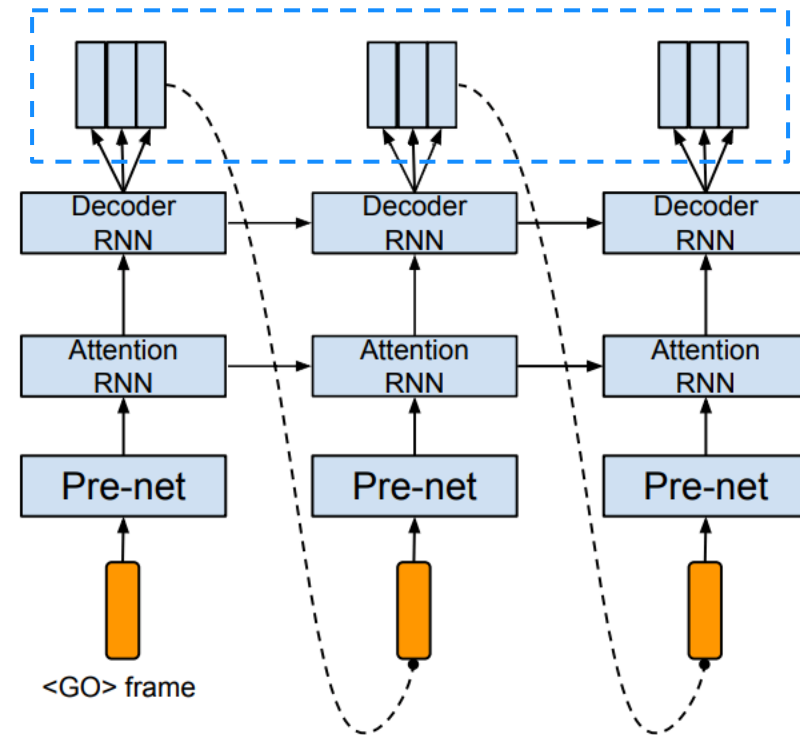
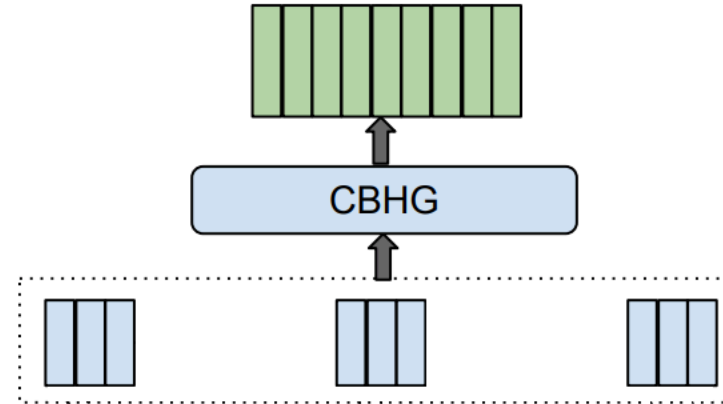
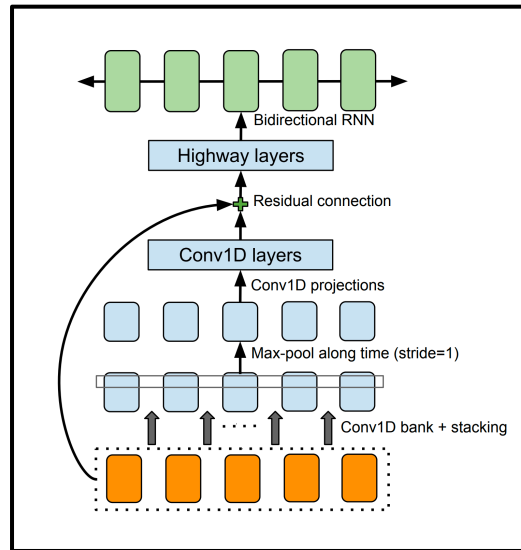
스팩트로그램들

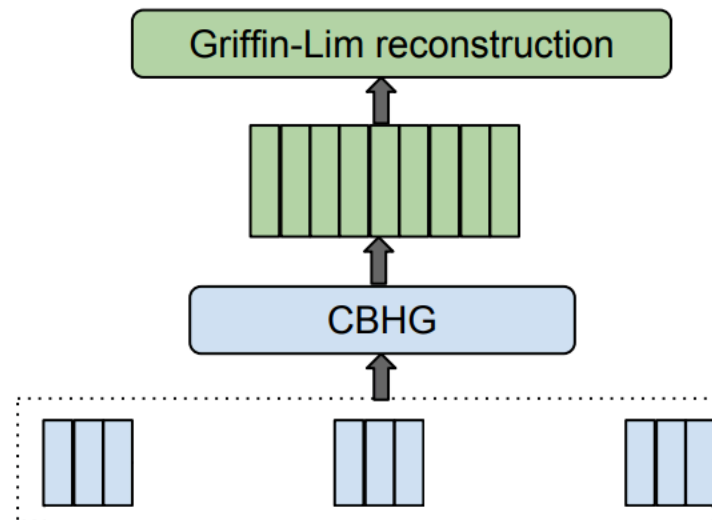
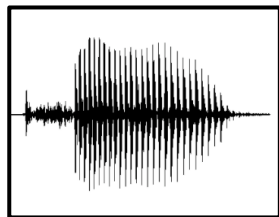
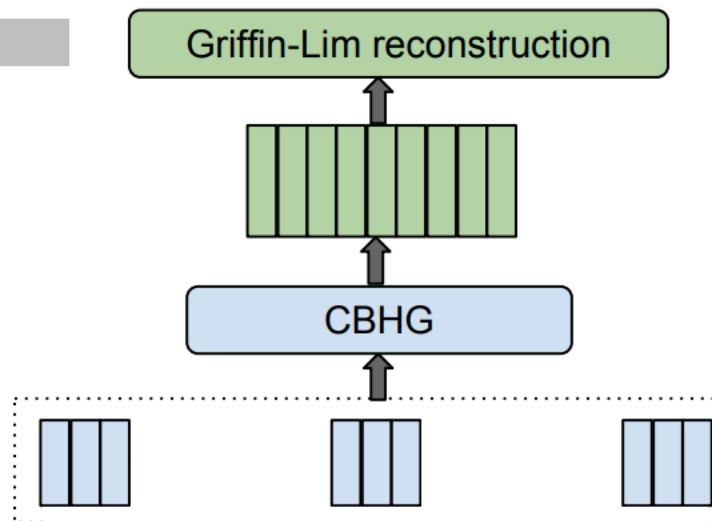
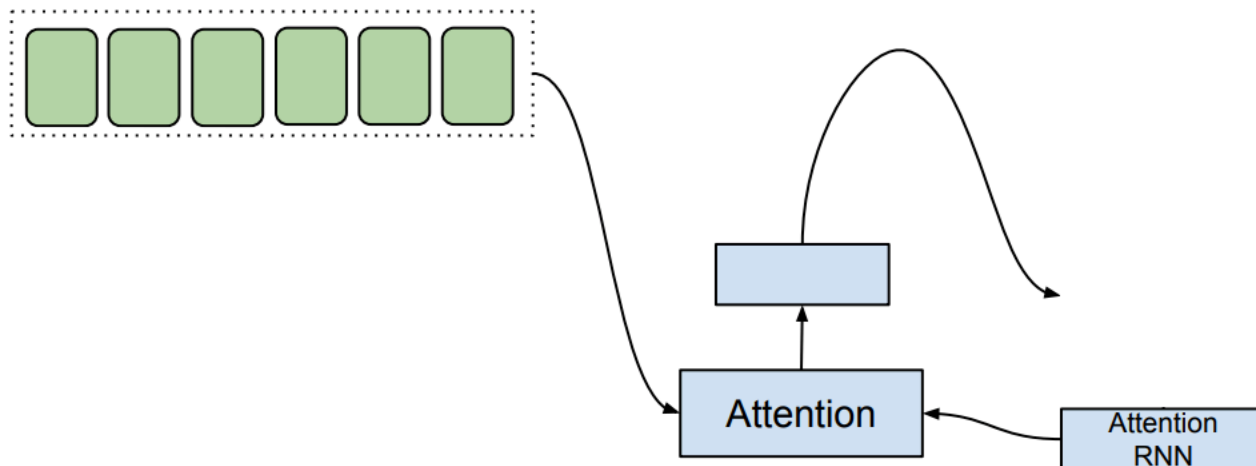
4. Vocoder

스펙트로그램들을

Decoder RNN
Decoder RNN
Decoder RNN
Attention RNN
Attention RNN
Attention RNN
Pre-net
Pre-net
Pre-net
<GO> frame

CBHG를 거치고

Bidirectional RNN
Highway layers
Residual connection
Conv1D layers
Conv1D projections
Max-pool along time (stride=1)
Conv1D bank + stacking

CBHG

스팩트로그램을 음성으로 만들어주는 한가지 알고리즘

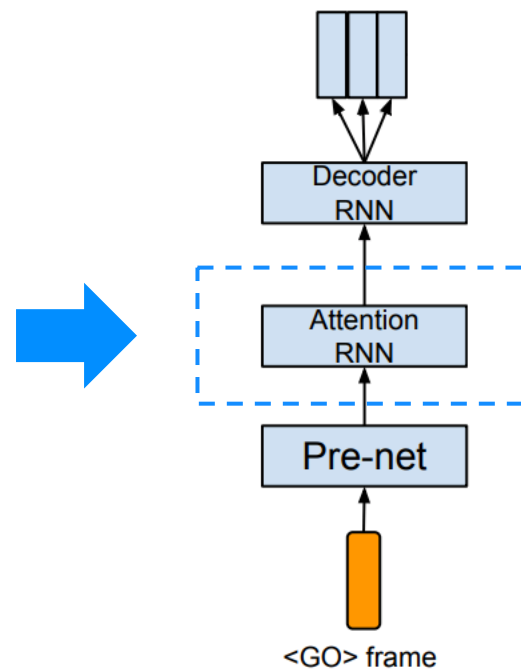
Griffin-Lim을 거쳐서

Griffin-Lim reconstruction
CBHG
음성을 만듦

3. Attention

Attention은 Tacotron에서 가장 중요

Attention

Attention

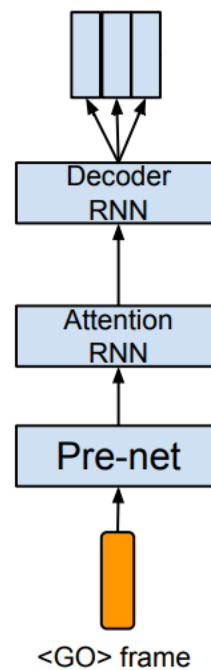
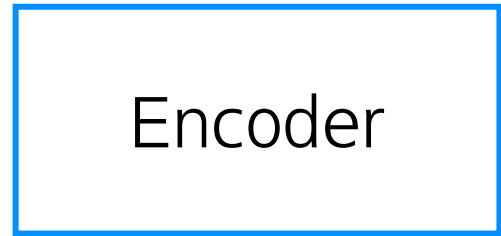
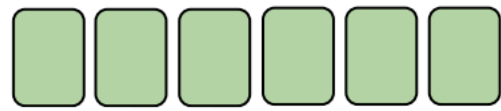
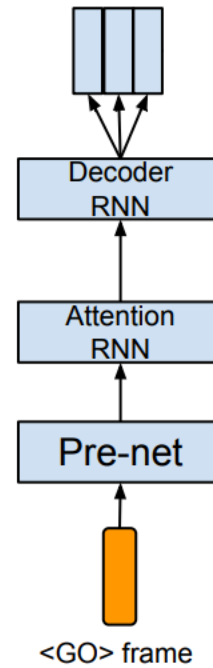
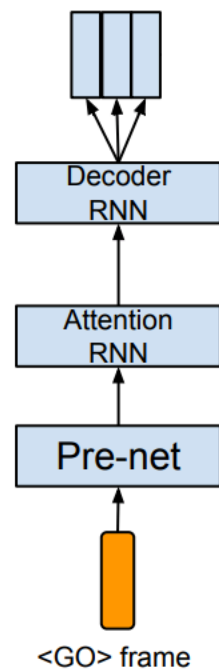
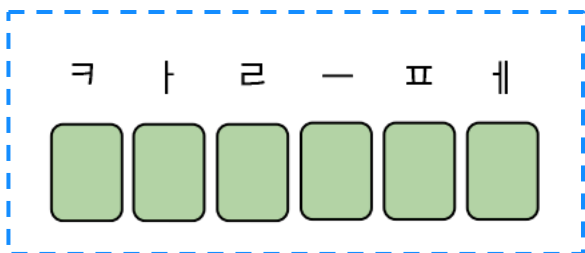
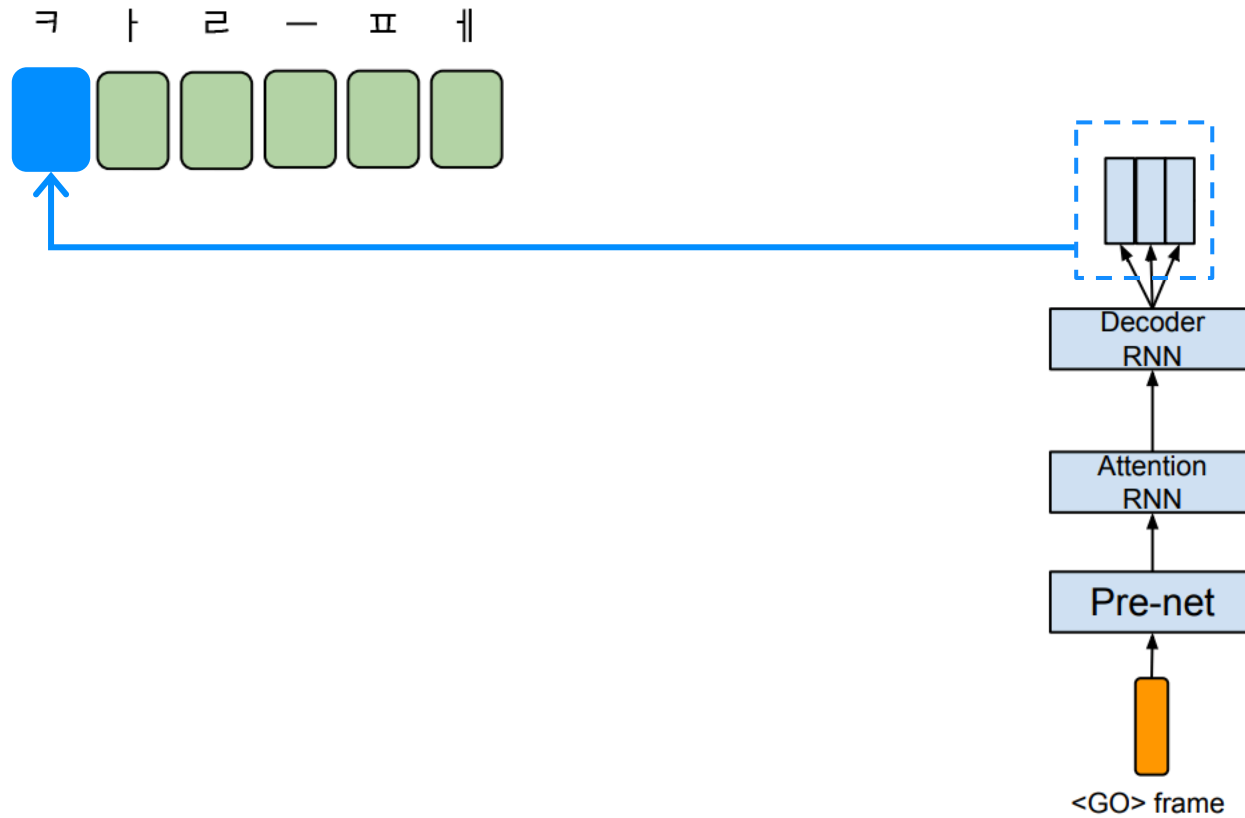
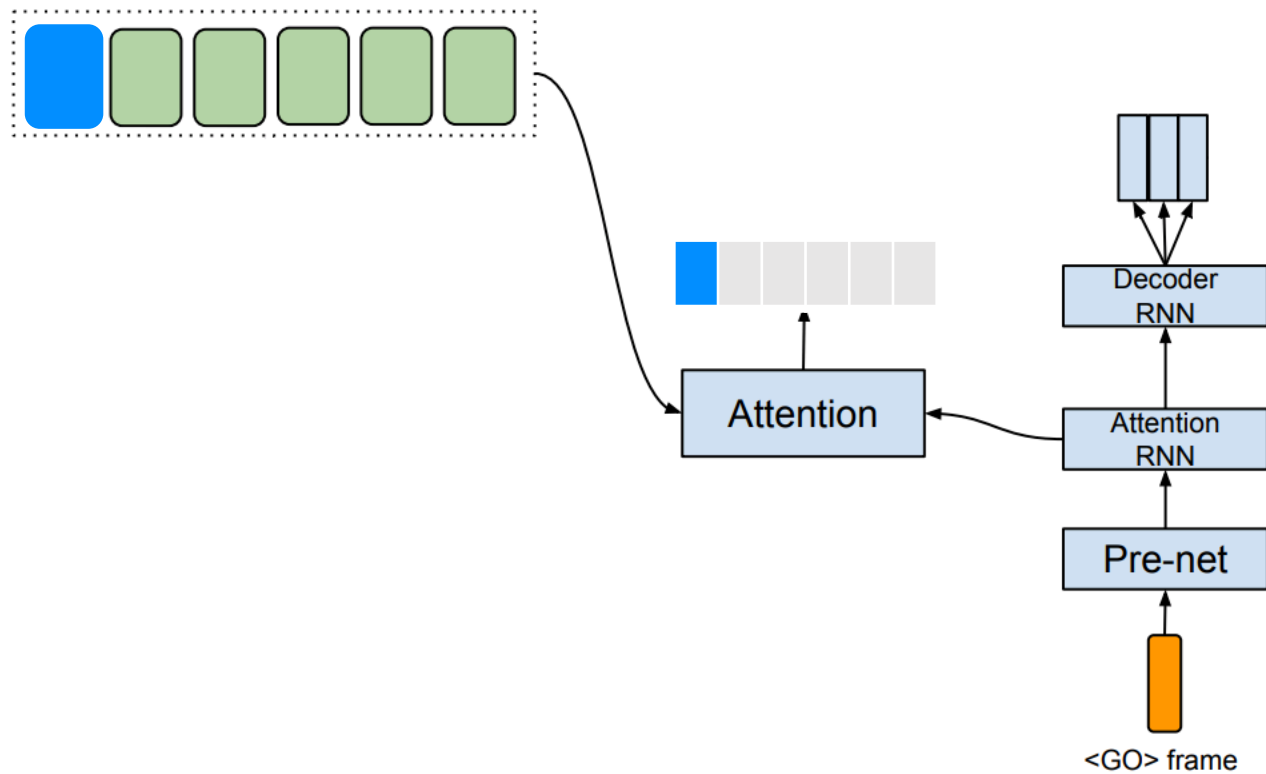
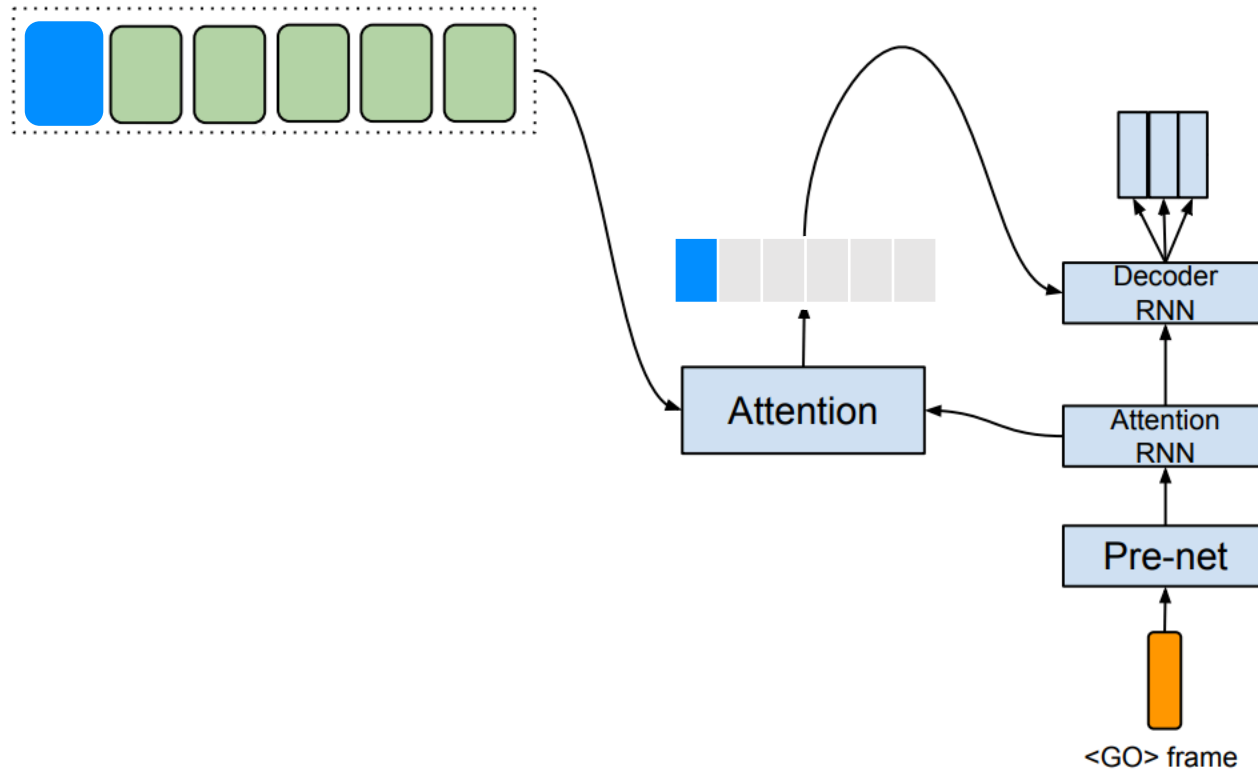
어디에 **집중**할 것인가?

Decoder
RNN
Attention
RNN
Pre-net
<GO> frame

텍스트 임베딩 : 텍스트를 잘 나타내는 숫자

텍스트 임베딩

ㅋ ㅏ ㄹ ㅡ ㅍ ㅔ

ㅋ ㅏ ㄹ ㅡ ㅍ ㅔ

Decoder RNN
Attention RNN
Pre-net
<GO> frame

첫번째 음성은 텍스트의 첫번째 단어에 집중

어디에 집중할 것인가를 계산하고

스펙트로그램을 만드는 RNN에 Attention을 전달

예시를 한번 보죠

텍스트 임베딩

당연히 우리나라 기름값도 내려갔지요.
ㄷ ㅏ ㅇ ㅇ ㅕ ㄴ ㅎ ㅣ ㅇ ㅜ ㄹ ㅣ ㄴ ㅏ ㄹ ㅏ ㄱ ㅣ ㄹ ㅡ ㅁ ㄱ ㅏ ㅄ ㄷ ㅗ ㄴ ㅐ ㄹ ㅕ ㄱ ㅏ ㅆ ㅈ ㅣ ㅇ ㅛ . ~
<Go>
가장 첫번째 단어에 집중

당연히 우리나라 기름값도 내려갔지요.

ㄷ ㅏ ㅇ ㅇ ㅕ ㄴ ㅎ ㅣ ㅇ ㅜ ㄹ ㅣ ㄴ ㅏ ㄹ ㅏ ㄱ ㅣ ㄹ ㅡ ㅁ ㄱ ㅏ ㅄ ㄷ ㅗ ㄴ ㅐ ㄹ ㅕ ㄱ ㅏ ㅆ ㅈ ㅣ ㅇ ㅛ . ~

Decoder RNN

당연히 우리나라 기름값도 내려갔지요.

ㄷ ㅏ ㅇ ㅇ ㅕ ㄴ ㅎ ㅣ ㅇ ㅜ ㄹ ㅣ ㄴ ㅏ ㄹ ㅏ ㄱ ㅣ ㄹ ㅡ ㅁ ㄱ ㅏ ㅄ ㄷ ㅗ ㄴ ㅐ ㄹ ㅕ ㄱ ㅏ ㅆ ㅈ ㅣ ㅇ ㅛ . ~

Decoder RNN

당연히 우리나라 기름값도 내려갔지요.

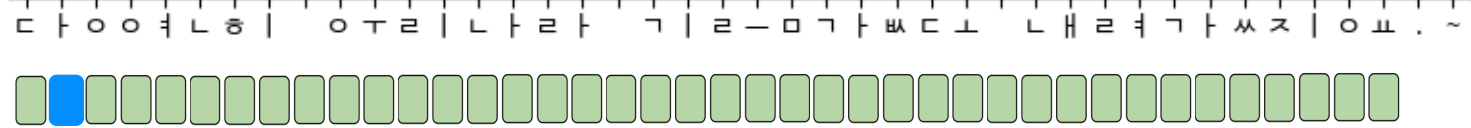
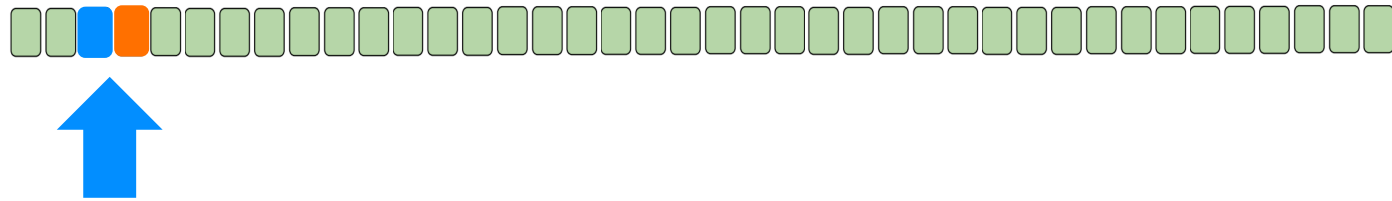
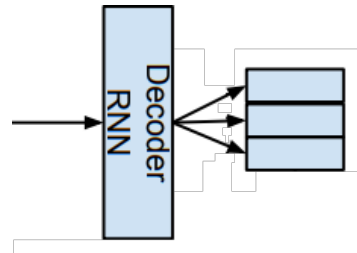
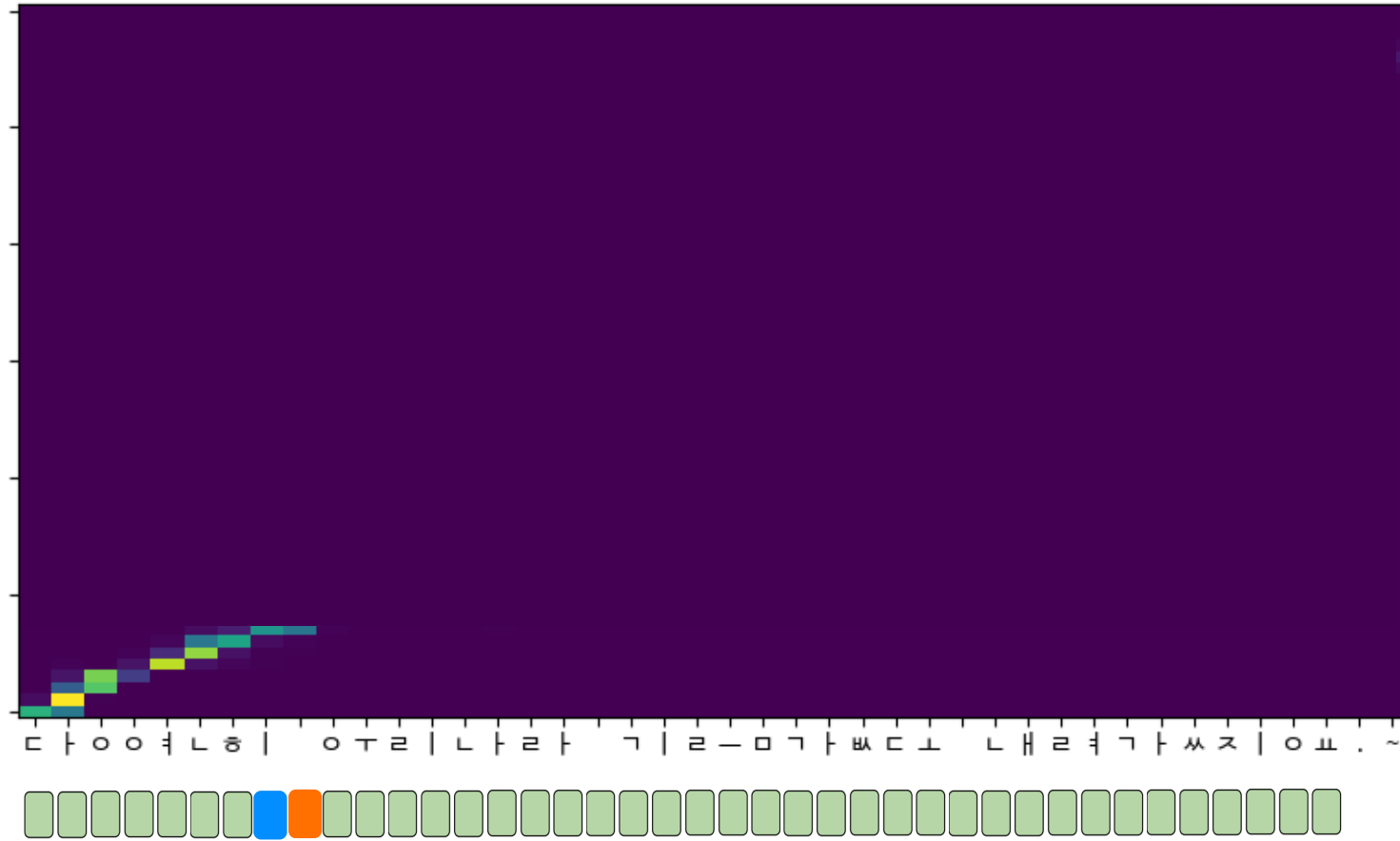
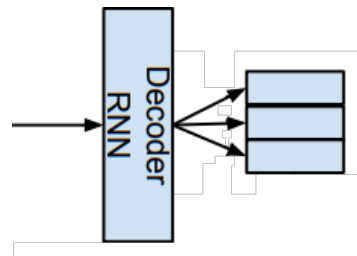
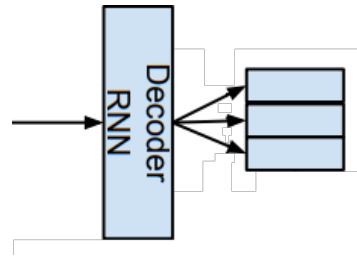
Decoder RNN

ㄷ ㅏ ㅇ ㅇ ㅕ ㄴ ㅎ ㅣ ㅇ ㅜ ㄹ ㅣ ㄴ ㅏ ㄹ ㅏ ㄱ ㅣ ㄹ ㅡ ㅁ ㄱ ㅏ ㅄ ㄷ ㅗ ㄴ ㅐ ㄹ ㅕ ㄱ ㅏ ㅆ ㅈ ㅣ ㅇ ㅛ . ~

한 단어 완성!

당연히 우리나라 기름값도 내려갔지요.

Decoder RNN

ㄷ ㅏ ㅇ ㅇ ㅕ ㄴ ㅎ ㅣ ㅇ ㅜ ㄹ ㅣ ㄴ ㅏ ㄹ ㅏ ㄱ ㅣ ㄹ ㅡ ㅁ ㄱ ㅏ ㅄ ㄷ ㅗ ㄴ ㅐ ㄹ ㅕ ㄱ ㅏ ㅆ ㅈ ㅣ ㅇ ㅛ . ~

공백을 보고 몇 초간 쉼

당연히 우리나라 기름값도 내려갔지요.

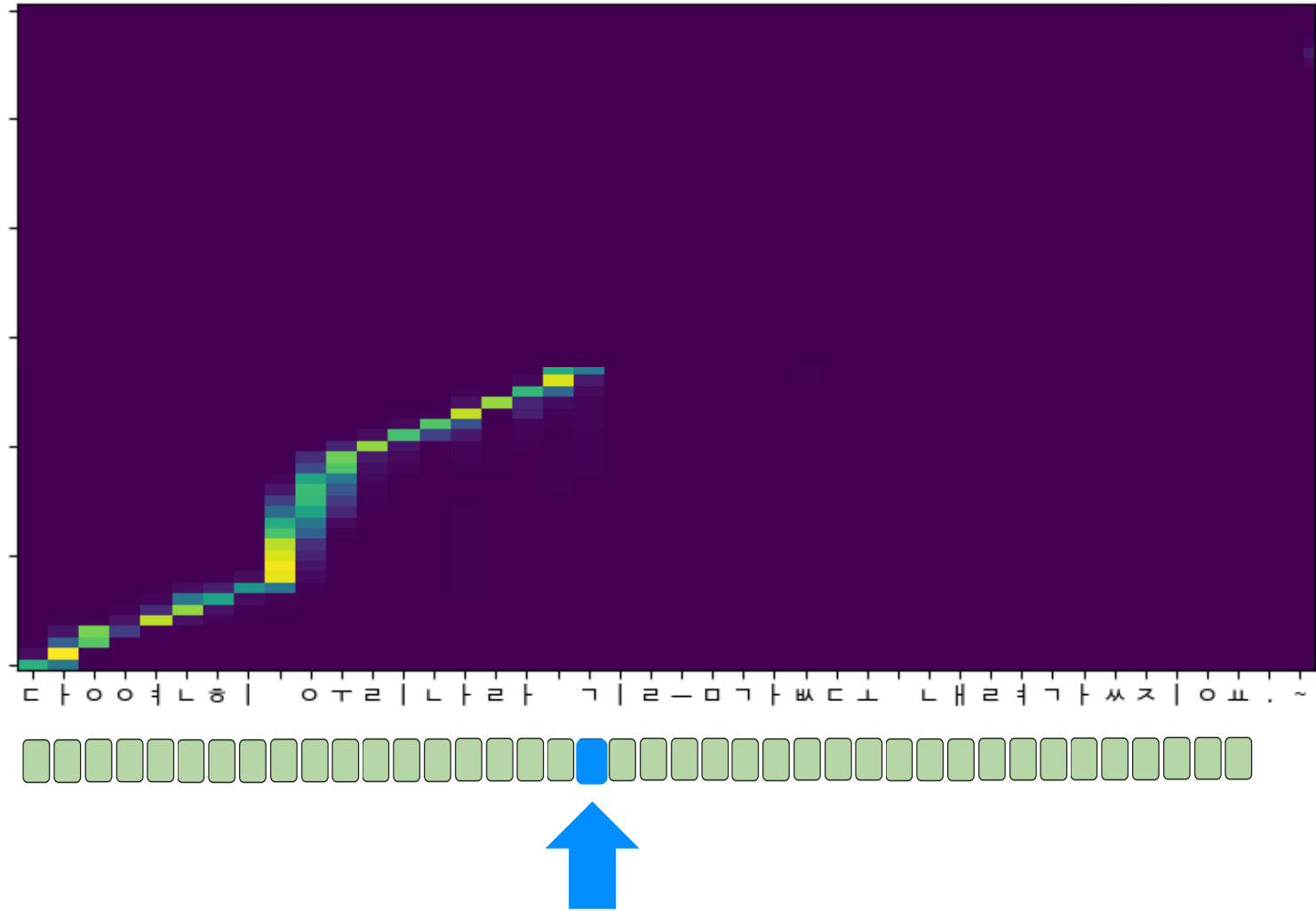
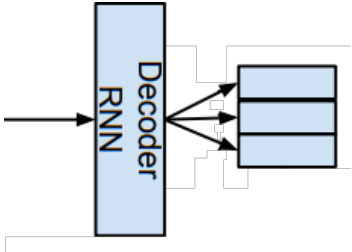
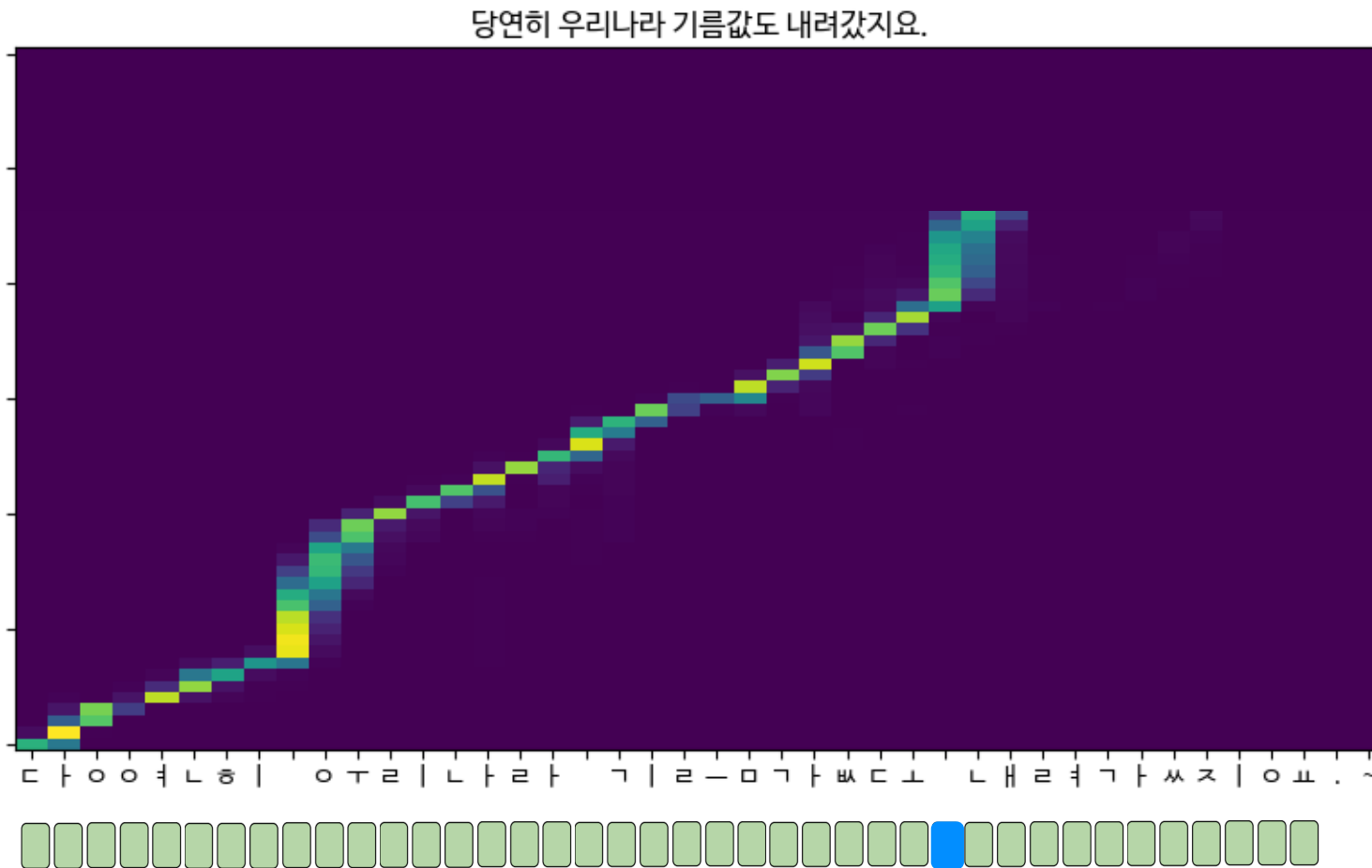
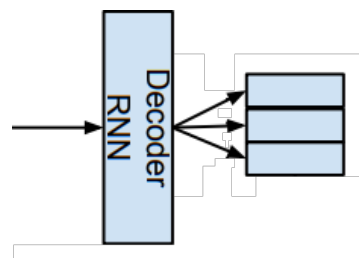
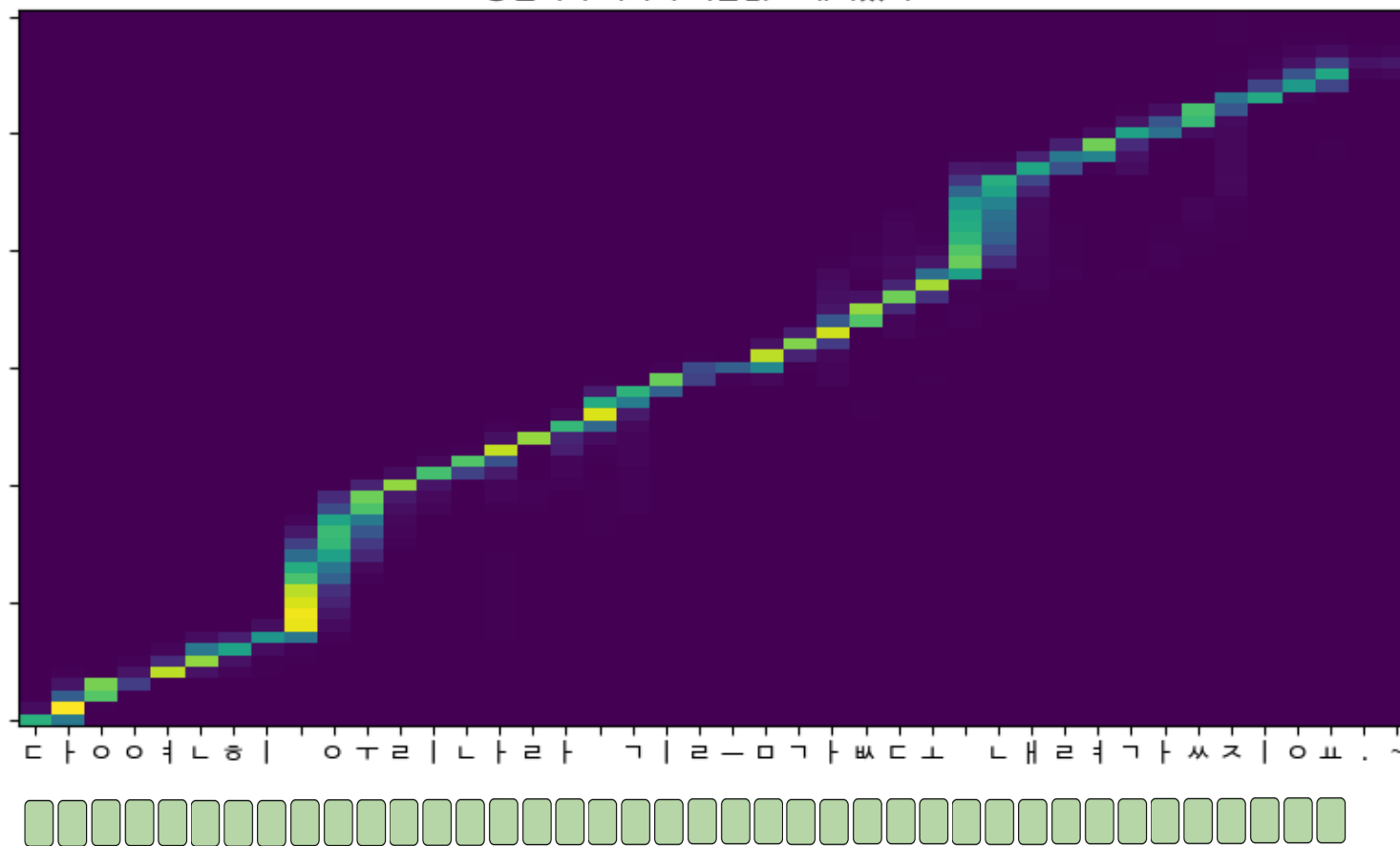
ㄷ ㅏ ㅇ ㅇ ㅕ ㄴ ㅎ ㅣ ㅇ ㅜ ㄹ ㅣ ㄴ ㅏ ㄹ ㅏ ㄱ ㅣ ㄹ ㅡ ㅁ ㄱ ㅏ ㅄ ㄷ ㅗ ㄴ ㅐ ㄹ ㅕ ㄱ ㅏ ㅆ ㅈ ㅣ ㅇ ㅛ . ~

Decoder RNN

Decoder RNN
당연히 우리나라 기름값도 내려갔지요.
ㄷ ㅏ ㅇ ㅇ ㅕ ㄴ ㅎ ㅣ ㅇ ㅜ ㄹ ㅣ ㄴ ㅏ ㄹ ㅏ ㄱ ㅣ ㄹ ㅡ ㅁ ㄱ ㅏ ㅄ ㄷ ㅗ ㄴ ㅐ ㄹ ㅕ ㄱ ㅏ ㅆ ㅈ ㅣ ㅇ ㅛ . ~

또 다른 공백

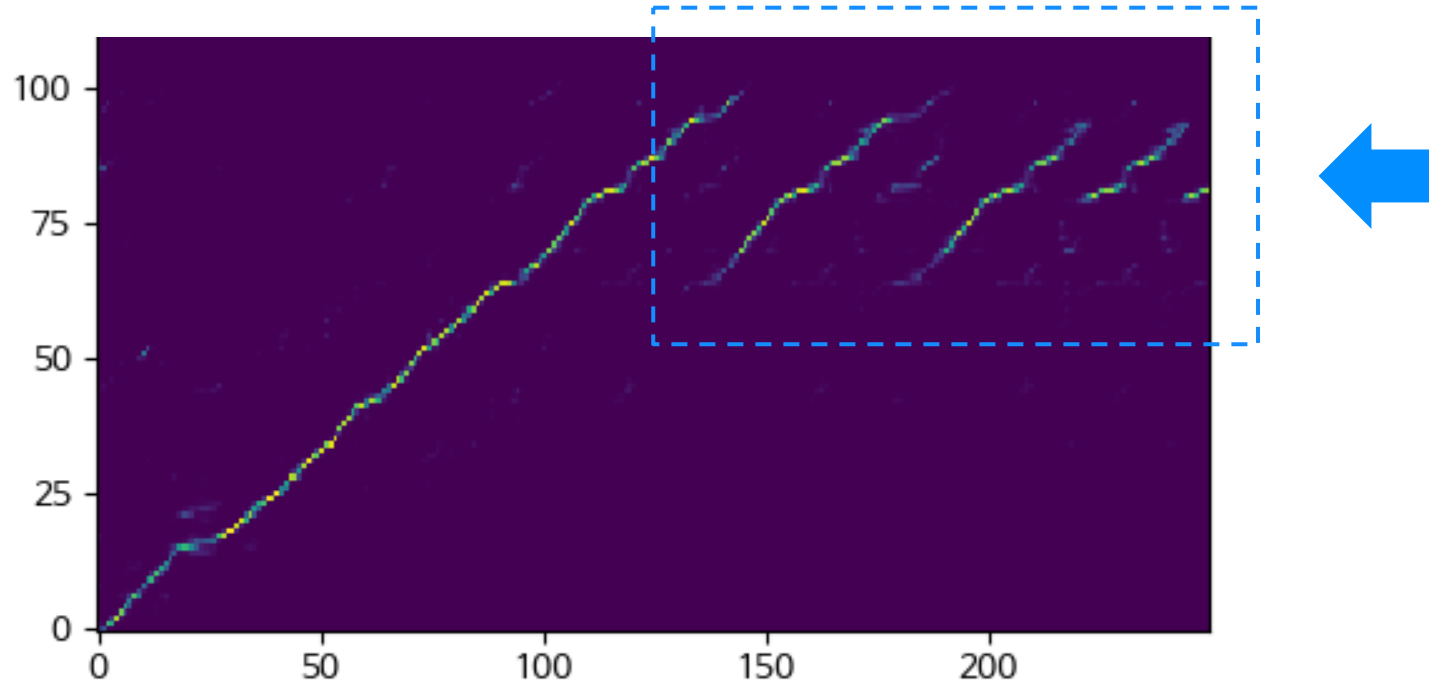
당연히 우리나라 기름값도 내려갔지요.

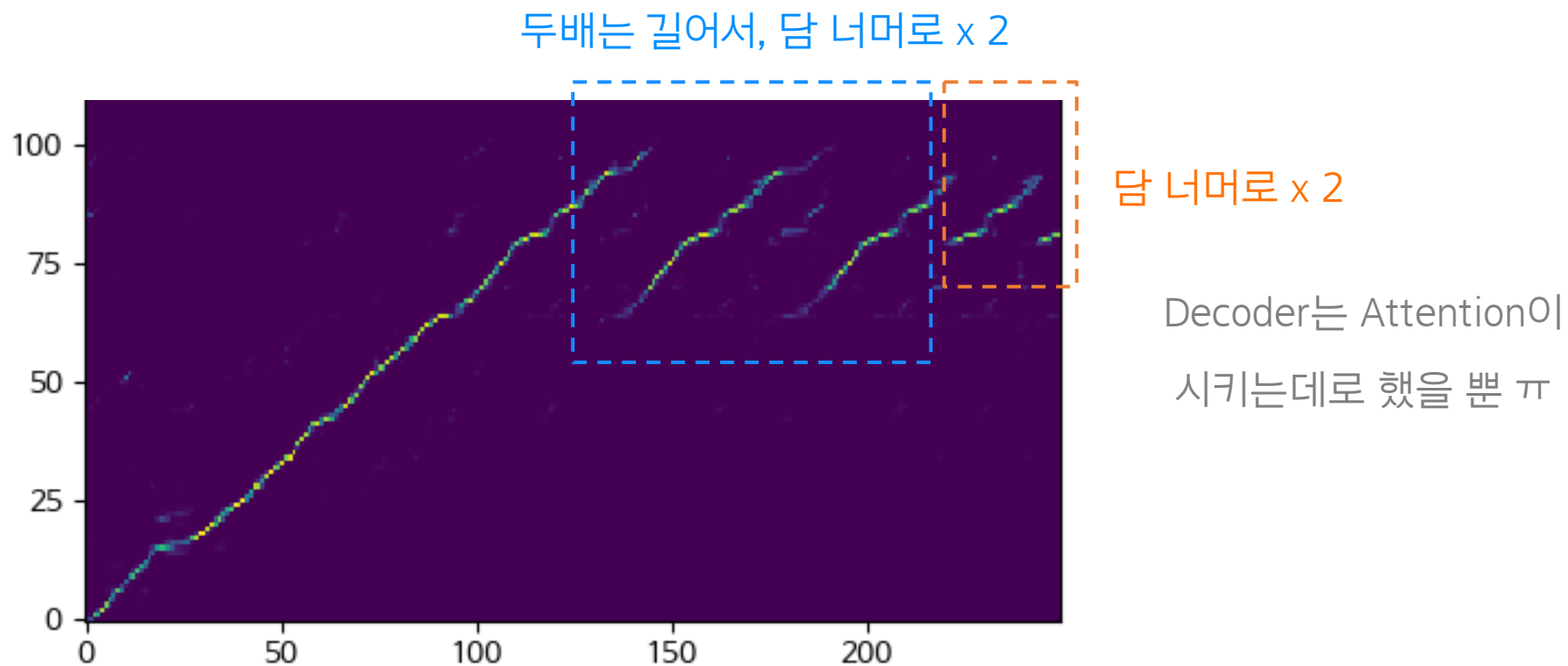
ㄷ ㅏ ㅇ ㅇ ㅕ ㄴ ㅎ ㅣ ㅇ ㅜ ㄹ ㅣ ㄴ ㅏ ㄹ ㅏ ㄱ ㅣ ㄹ ㅡ ㅁ ㄱ ㅏ ㅄ ㄷ ㅗ ㄴ ㅐ ㄹ ㅕ ㄱ ㅏ ㅆ ㅈ ㅣ ㅇ ㅛ . ~

Attention이 중요한 이유?

Generalization

일반화

음성 합성 모델의 핵심

학습하지 않았던 문장도 얼마나 잘 말할 수 있는가?

더즐리 부인은 마른 체구의 금발이었고, 목이 보통사람보다

두배는 길어서, 담 너머로 고개를 쭉 배고 이웃사람들을 몰래 훔쳐보는

그녀의 취미에 더 없이 제격이었다

더즐리 부인은 마른 체구의 금발이었고, 목이 보통사람보다

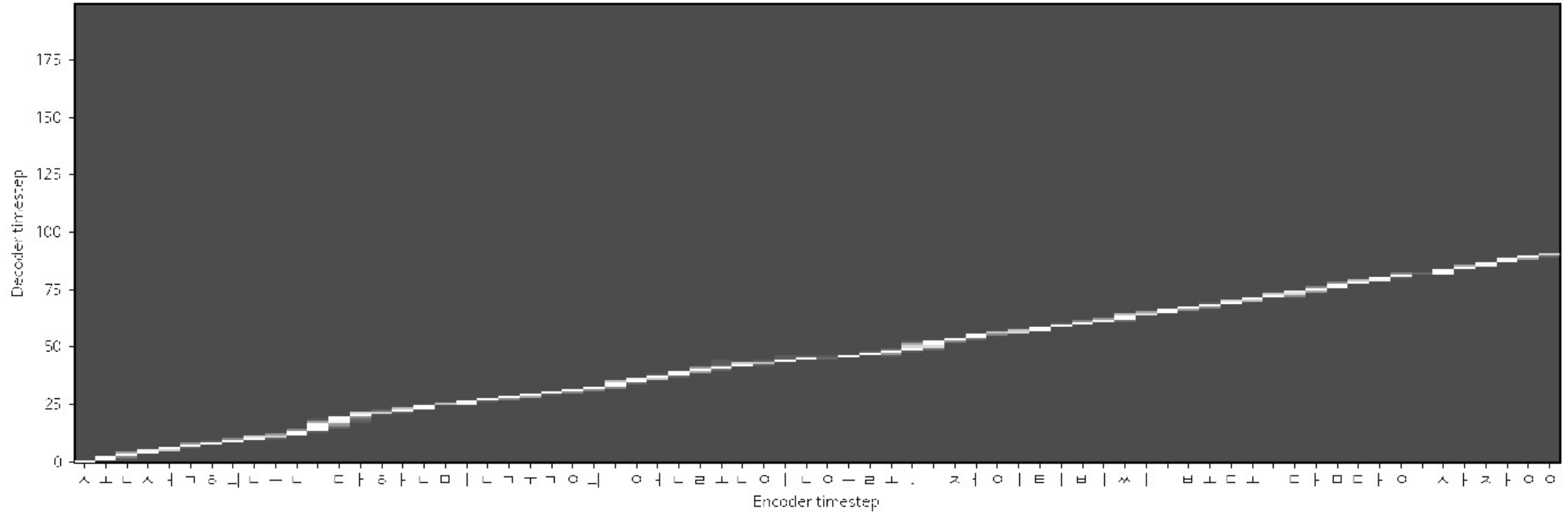
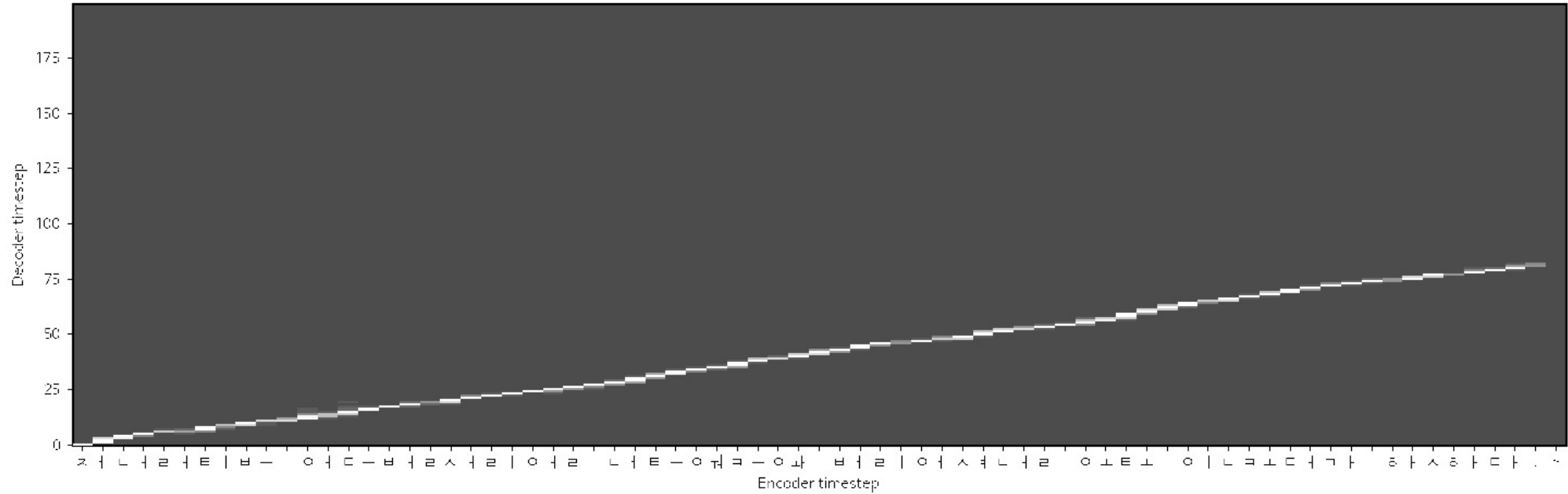
두배는 길어서, 담 너머로 고개를 쭉 배고 이웃사람들을 몰래 훔쳐보는

그녀의 취미에 더 없이 제격이었다

손석희는 대한민국의 언론인으로, JTBC 보도 담당 사장이다.

Decoder timestep

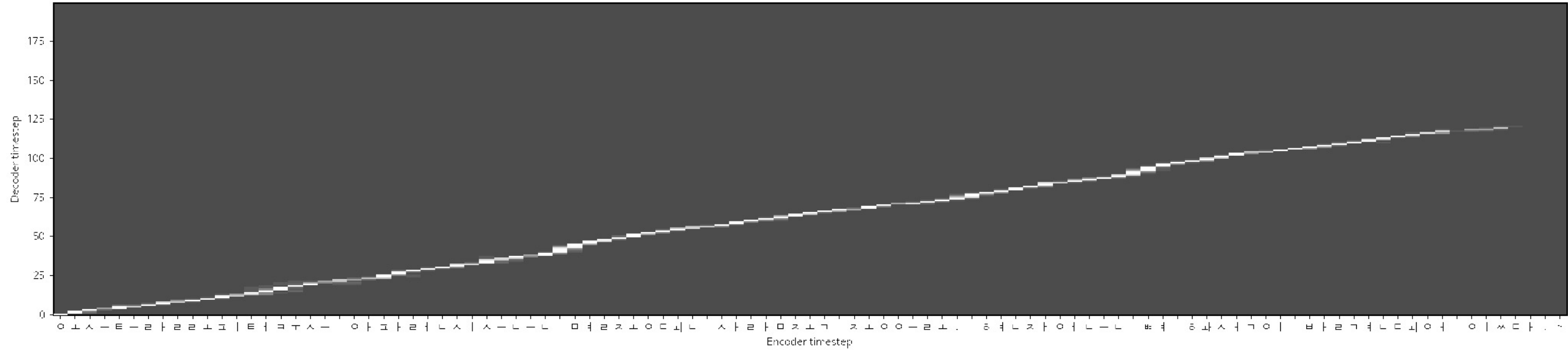
175
150
125
100
75
50
25
0

Encoder timestep

손석희는 대한민국의 언론인으로, JTBC 보도 담당 사장이다.

제너러티브 어드벌서리얼 네트워크와

베리에셔널 오토 인코더가 핫하다.

오스트랄로피테쿠스 아파렌시스는 멸종된 사람족 종으로, 현재에는 뼈 화석이 발견되어 있다.

Decoder timestep

Encoder timestep

오스트랄로피테쿠스 아파렌시스는 멸종된 사람족 종으로,

현재에는 뼈 화석이 발견되어 있다.

Attention의 학습 과정을 보면

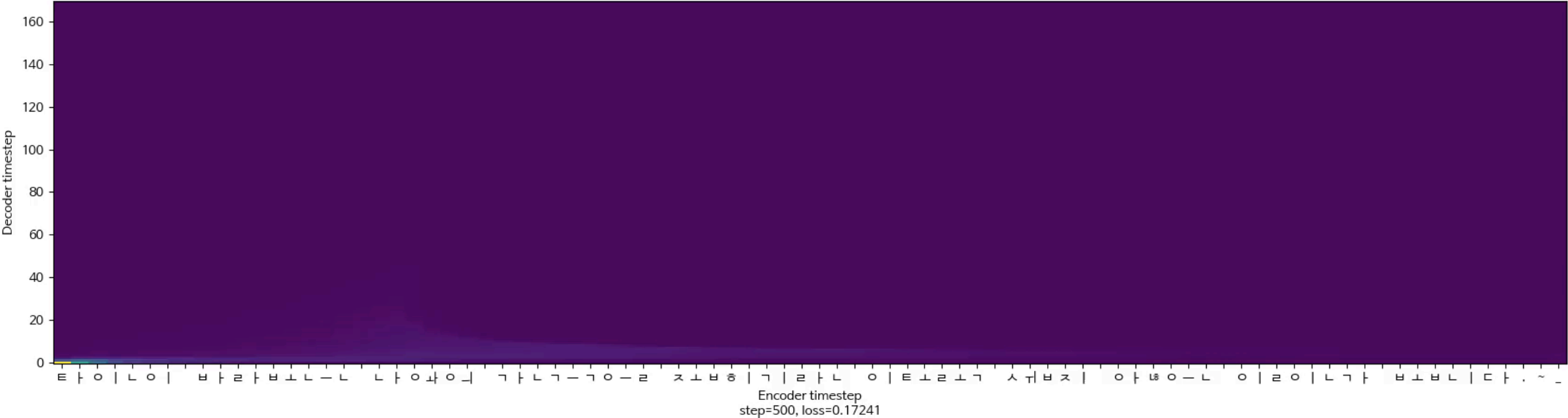
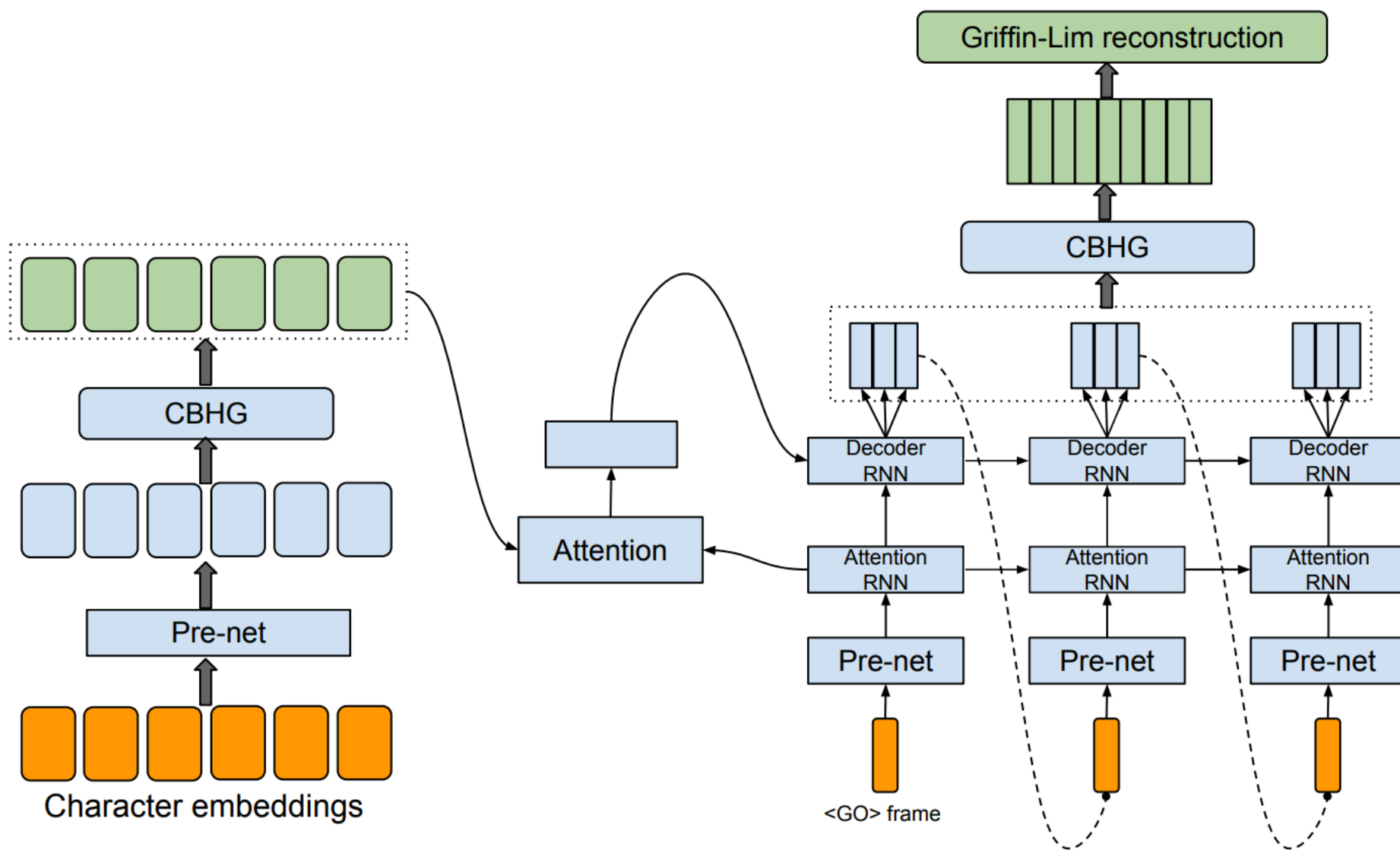
타인이 바라보는 나와의 간극을 좁히기란 이토록 쉽지 않은 일인가 봅니다.
Decoder timestep
160
140
120
100
80
60
40
20
0
Encoder timestep
step=500, loss=0.17241

Decoder timestep

Encoder timestep

step=500, loss=0.17241

네

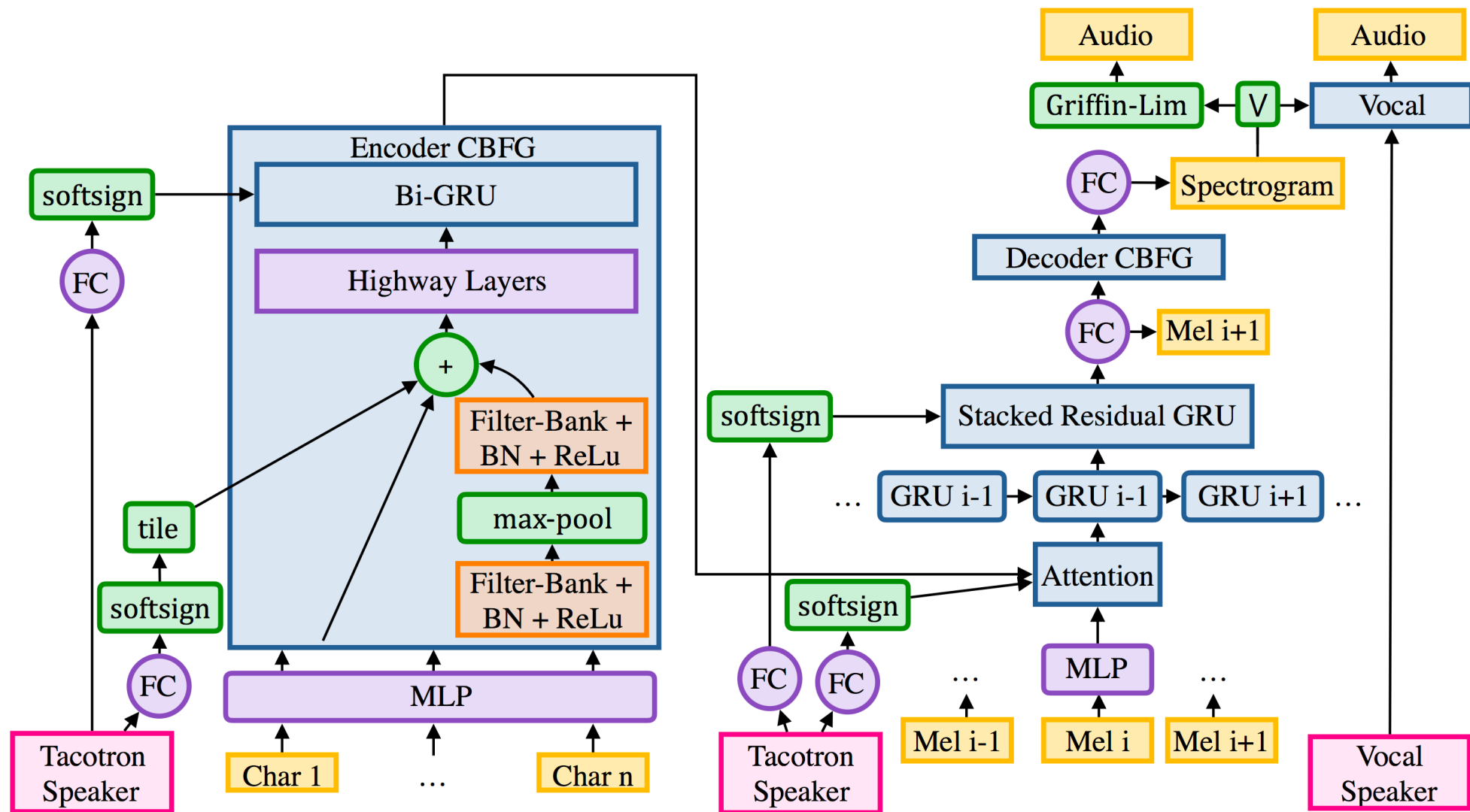
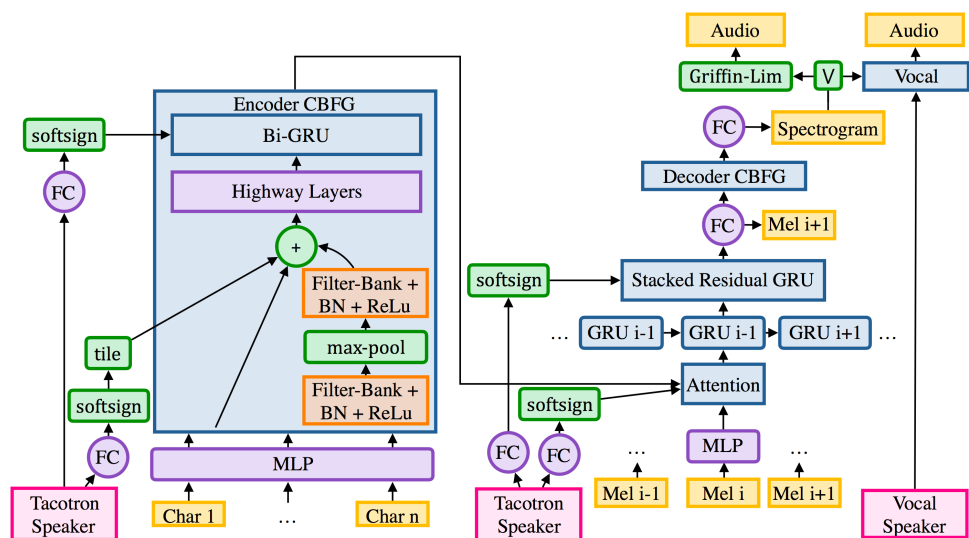
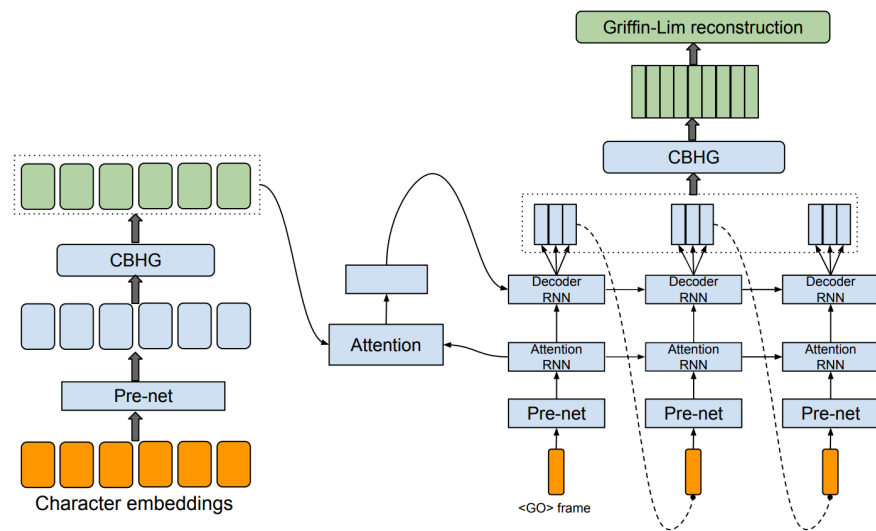
Wang, Yuxuan, et al. "Tacotron: A Fully End-to-End Text-To-Speech Synthesis Model." arXiv preprint arXiv:1703.10135 (2017).

Embedding

Bidirectional-RNN

Attention

2. Deep Voice 2

5달 전

Arik, Sercan, et al. "Deep Voice 2: Multi-Speaker Neural Text-to-Speech." arXiv preprint arXiv:1705.08947 (2017).

그냥 똑같은걸 다시 그린 것

핵심은

Tacotron을 Multi-Speaker 모델로

N명의 목소리를 하나의 모델로

GeForce GTX 980 Ti

Tacotron = 5G GPU

5명의 목소리를 만들고 싶다면?

5 × GeForce GTX 980 Ti

$5 \times$ Tacotron = 25 GB GPU

하지만 N-speaker Tacotron과 함께라면?

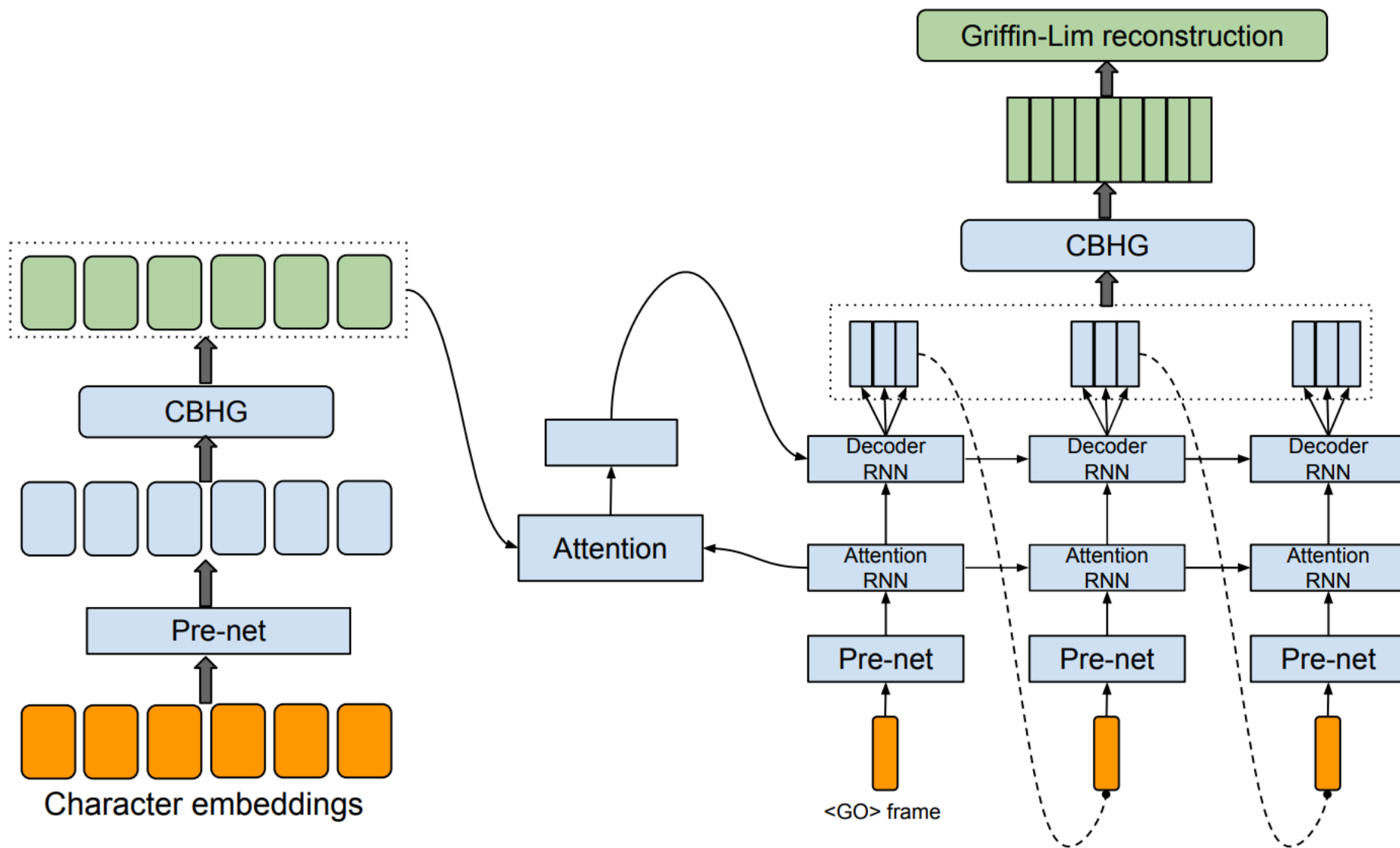
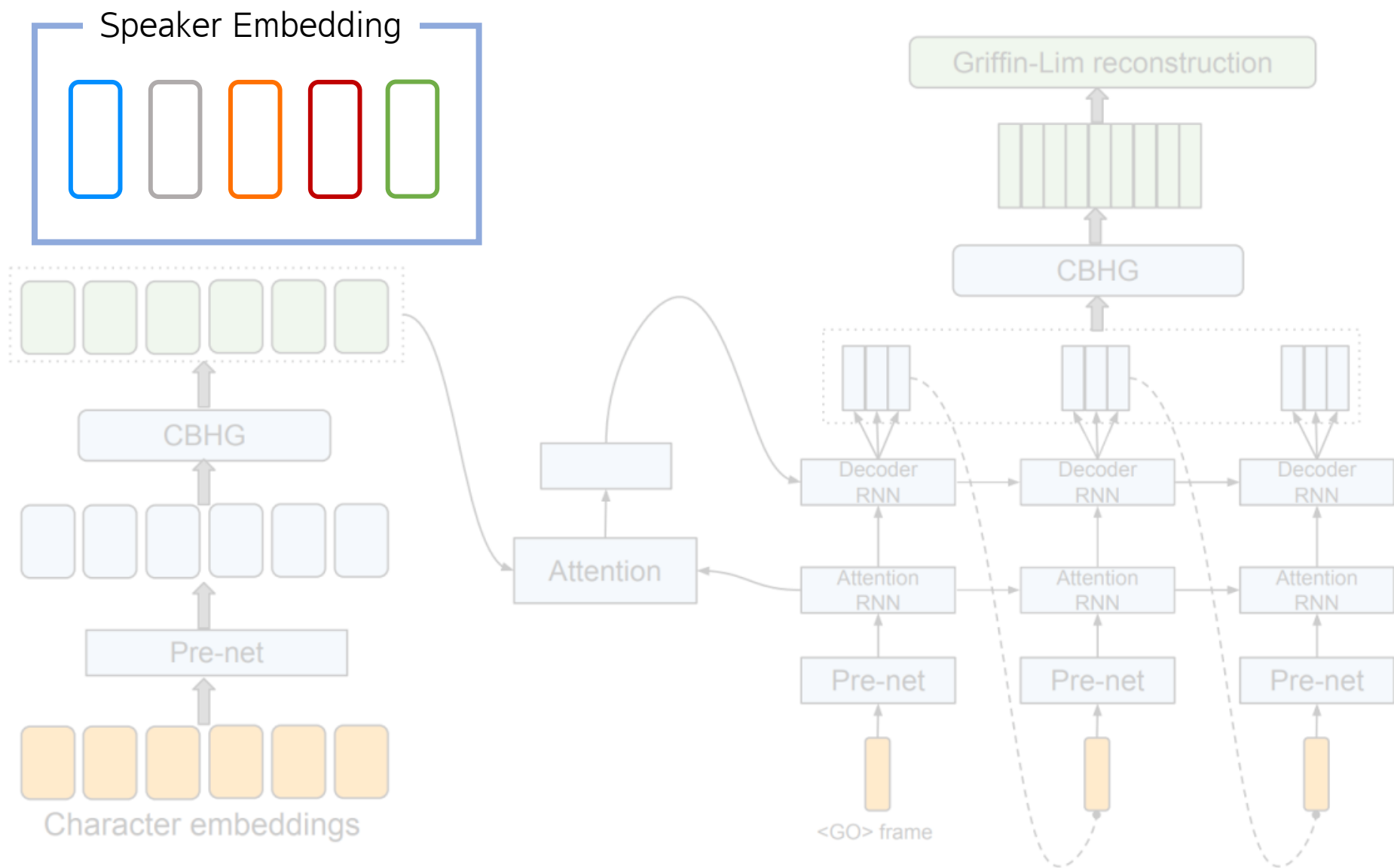
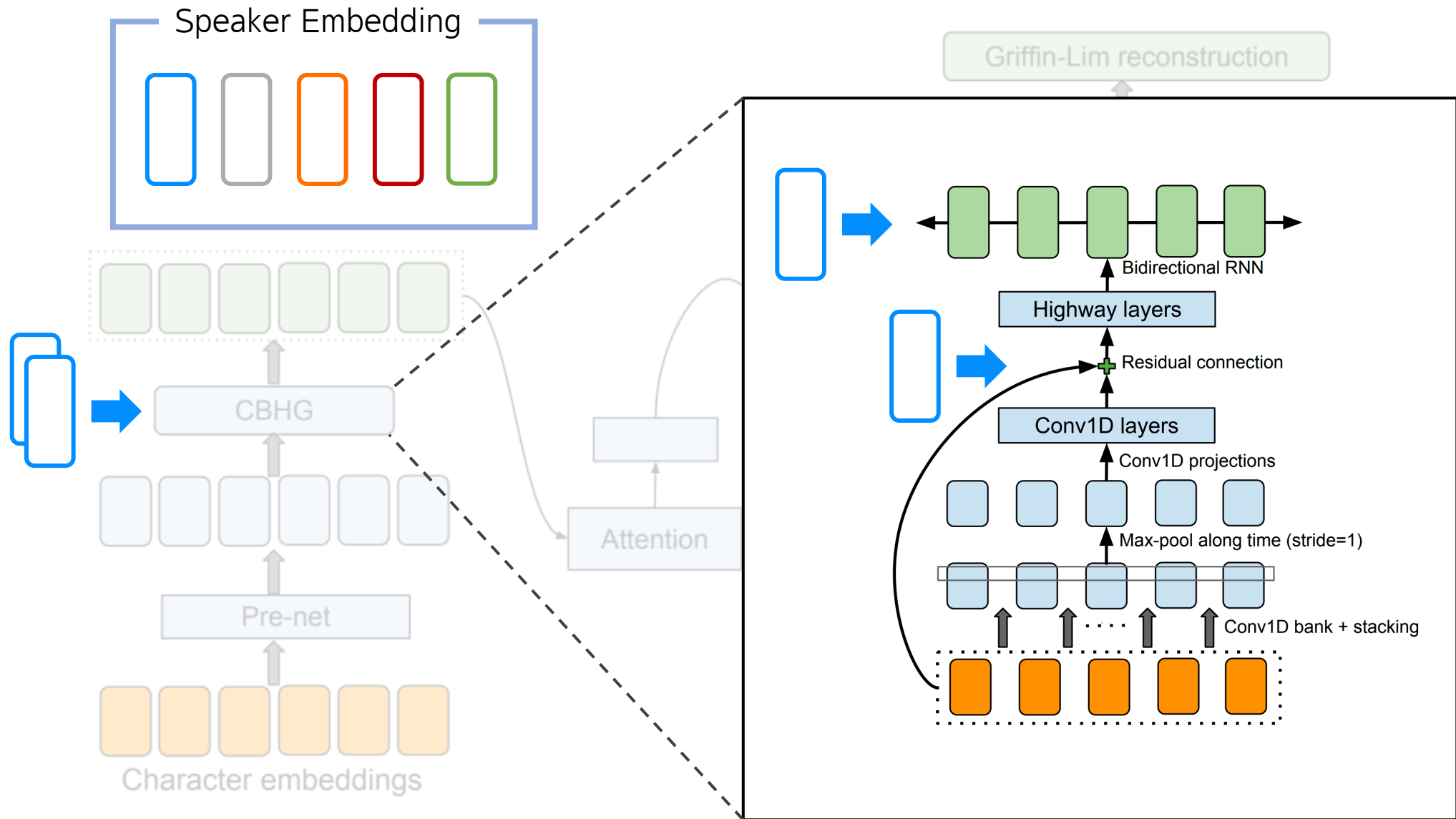
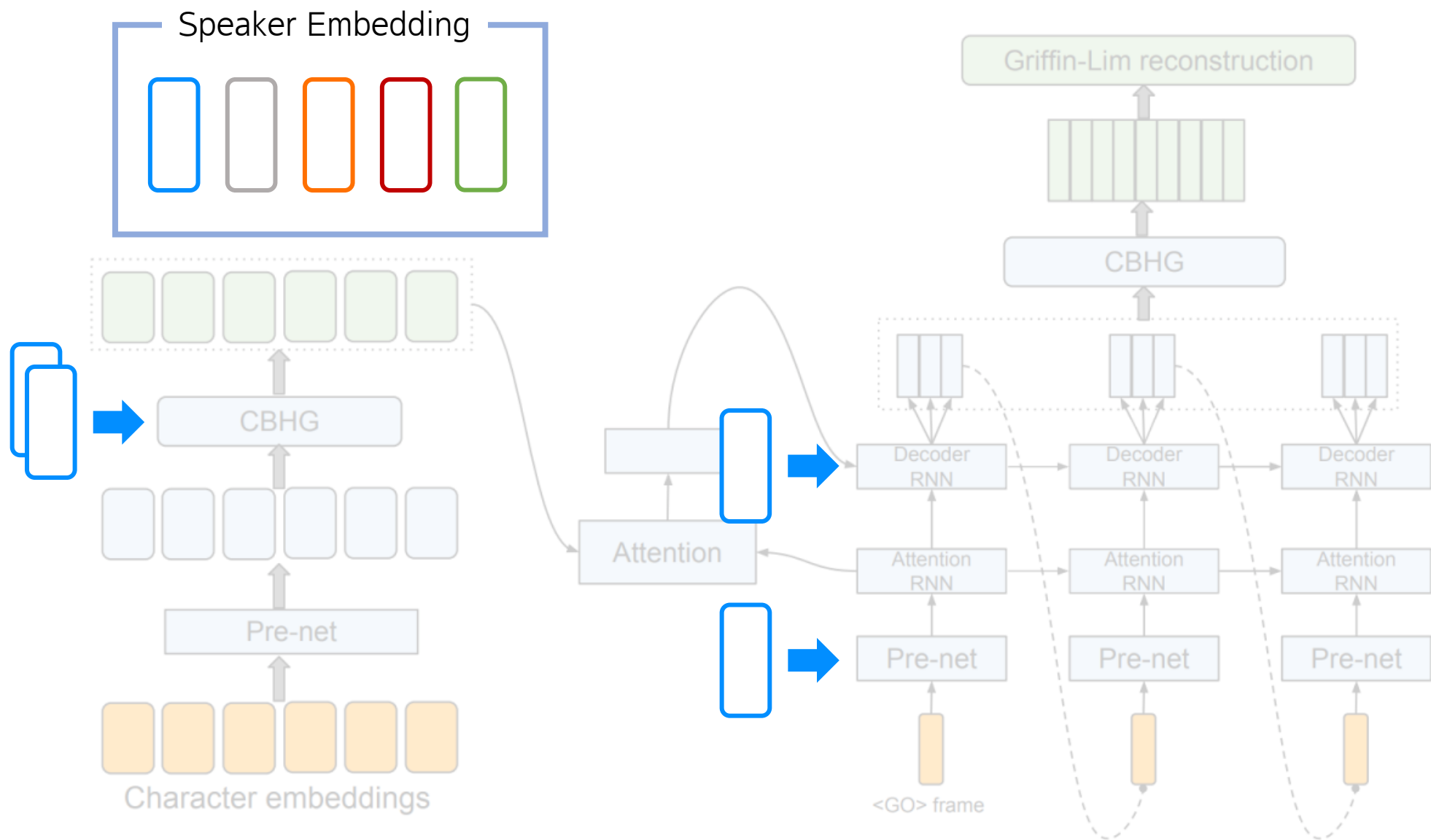
1 × GeForce GTX 980 Ti

5-speaker Tacotron ≅ 5G GPU

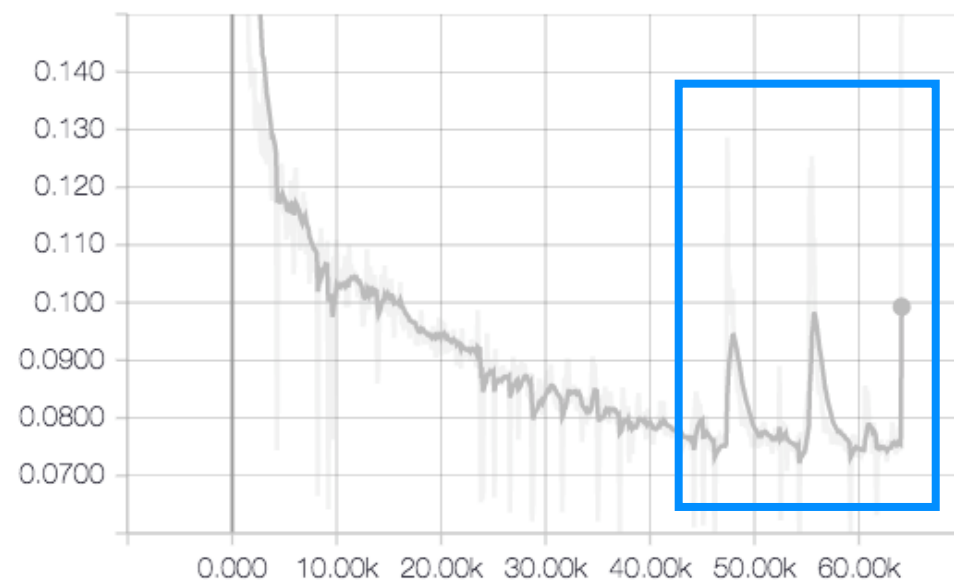
어떻게?

Speaker Embedding

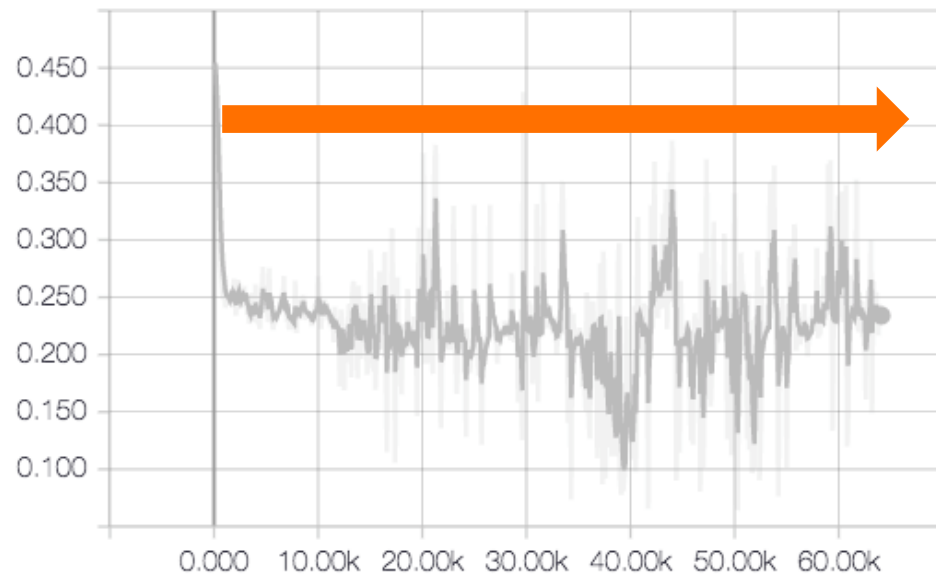
김태훈	유인나	손석희	네이버
0.7	0.3	0.1	0.9
0.5	0.0	0.0	0.5
0.9	0.7	0.3	0.1
0.5	0.5	0.0	0.1

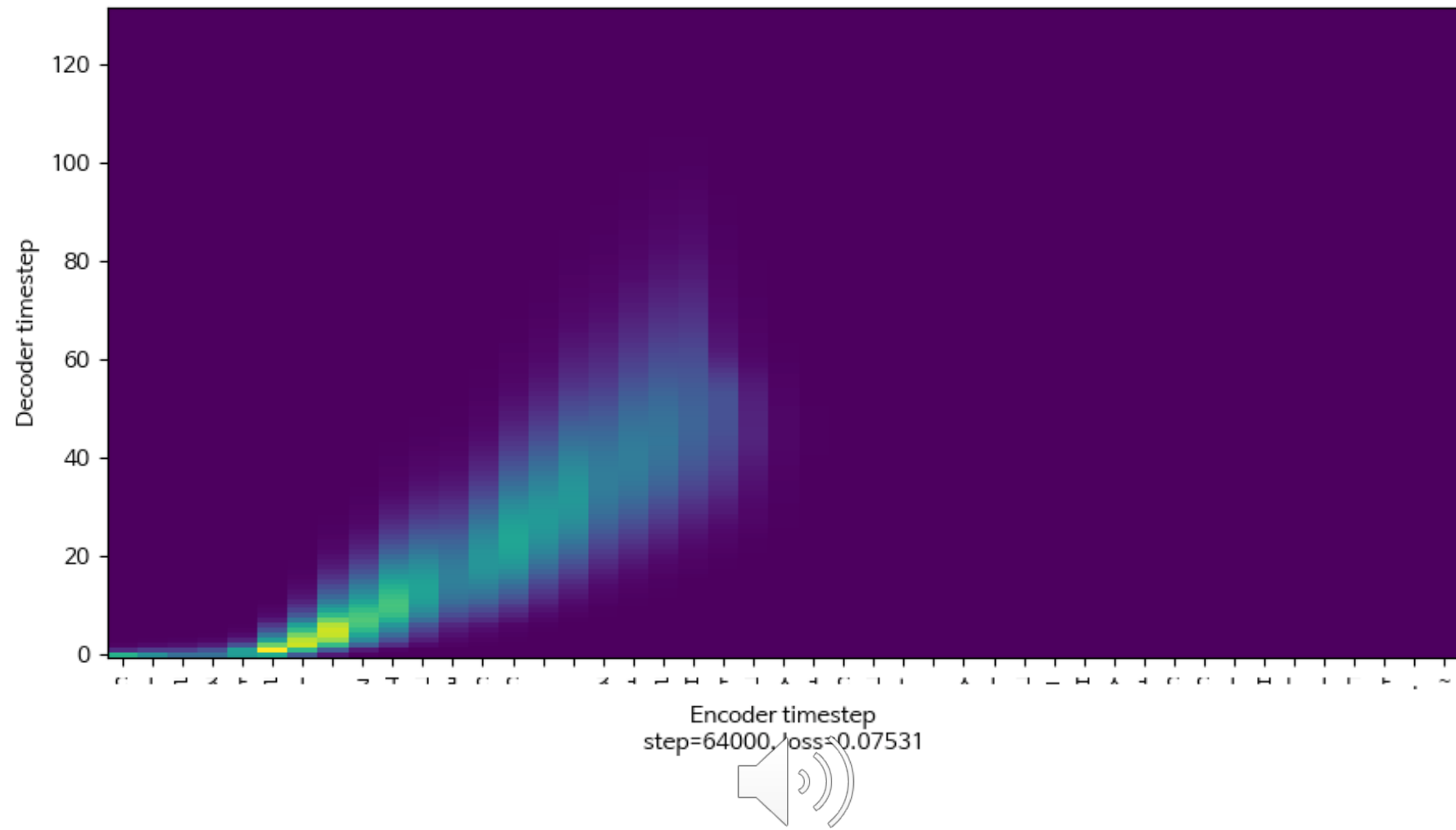
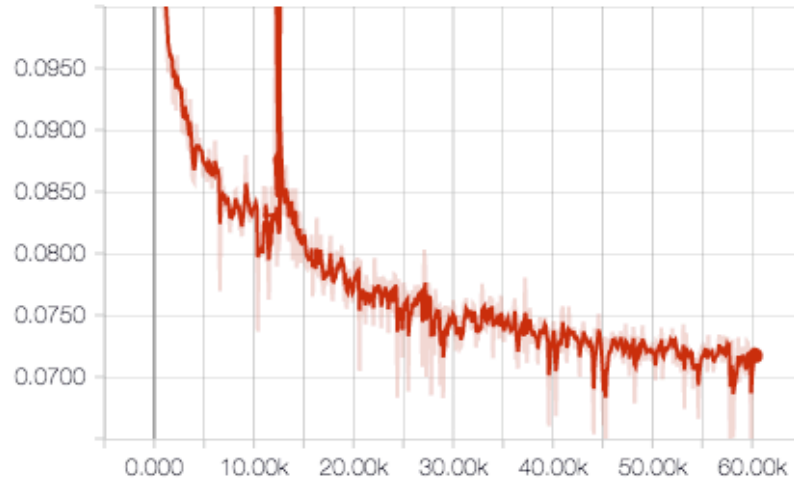
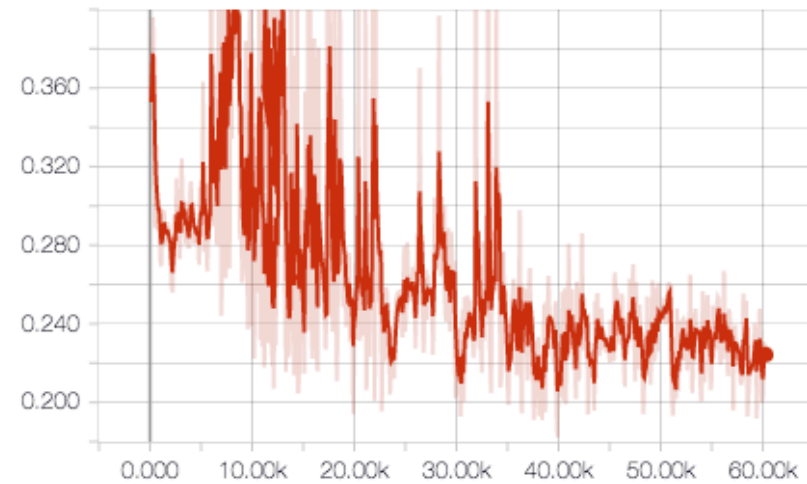
Griffin-Lim reconstruction
CBHG
CBHG
Pre-net
Character embeddings
Attention
Decoder RNN
Decoder RNN
Decoder RNN
Attention RNN
Attention RNN
Attention RNN
Pre-net
Pre-net
Pre-net
<GO> frame

또 다른 **장점**

집권 기간이 짧으셔서.. 2시간+

문재인 대통령 데이터 학습

model/stats/loss
0.140
0.130
0.120
0.110
0.100
0.0900
0.0800
0.0700
0.000
10.00k
20.00k
30.00k
40.00k
50.00k
60.00k
폭.발!

Train loss

test/loss_mel
0.450
0.400
0.350
0.300
0.250
0.200
0.150
0.100
0.000
10.00k
20.00k
30.00k
40.00k
50.00k
60.00k

Test loss

일자리 추경의 절박성과 시급성입니다.
Decoder timestep
Encoder timestep
step=64000, loss=0.07531

하지만 **다른 데이터**와 **함께** 학습하면?

Attention을 배워야 Test loss가 감소

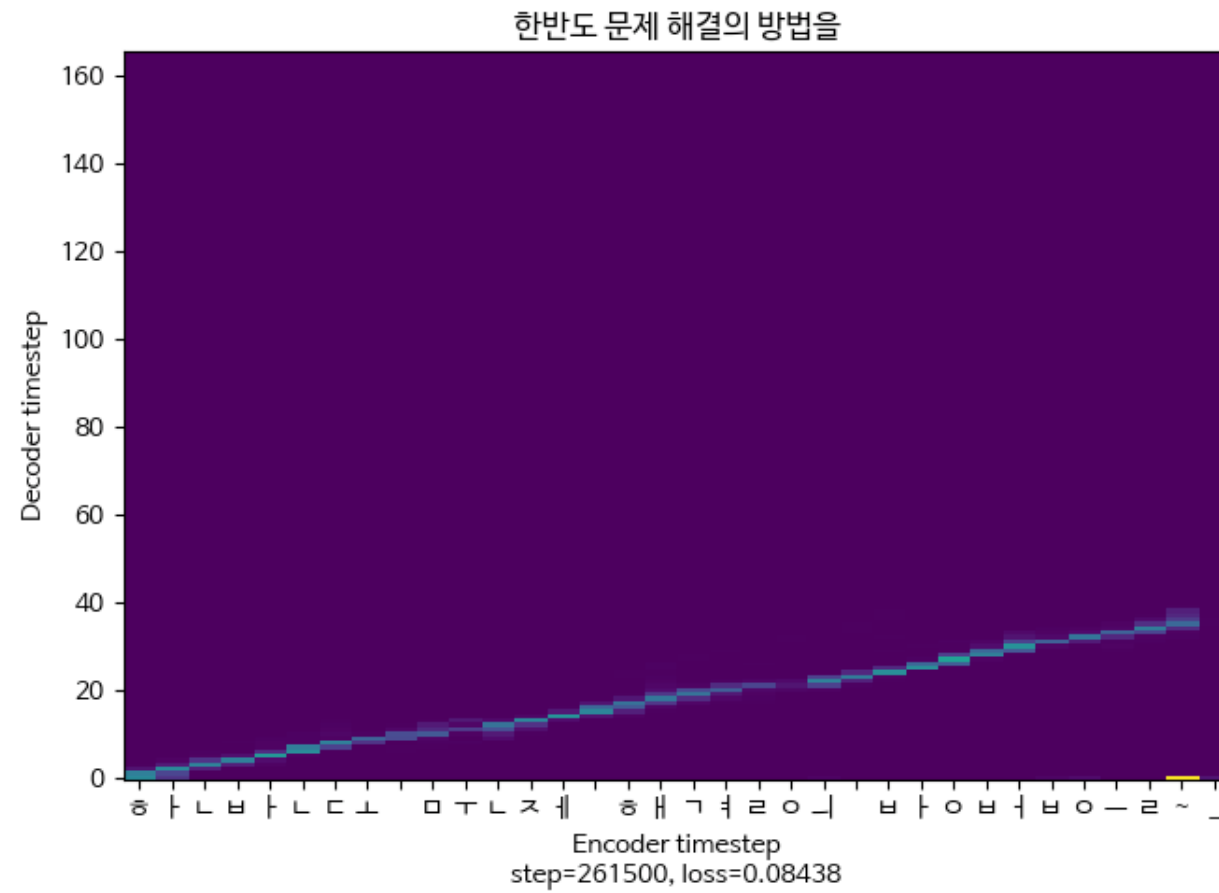
한반도 문제 해결의 방법을

Decoder timestep

160
140
120
100
80
60
40
20
0

ㅎ ㅏ ㄴ ㅂ ㅏ ㄴ ㄷ ㅗ ㅁ ㅜ ㄴ ㅈ ㅔ ㅎ ㅐ ㄱ ㅕ ㄹ ㅇ ㅢ ㅂ ㅏ ㅇ ㅂ ㅓ ㅂ ㅇ ㅡ ㄹ ~ _

Encoder timestep

step=261500, loss=0.08438

학습을 **도와준다**

완벽한 데이터 + 완벽하지 않은 데이터

Multi-speaker Attention

목소리 A

목소리 B

또한 그 통제가 무너졌을 때 다가올 어두움의 저편까지도

Decoder timestep

Encoder timestep

step=500, loss=0.17241

목소리 C

목소리 D

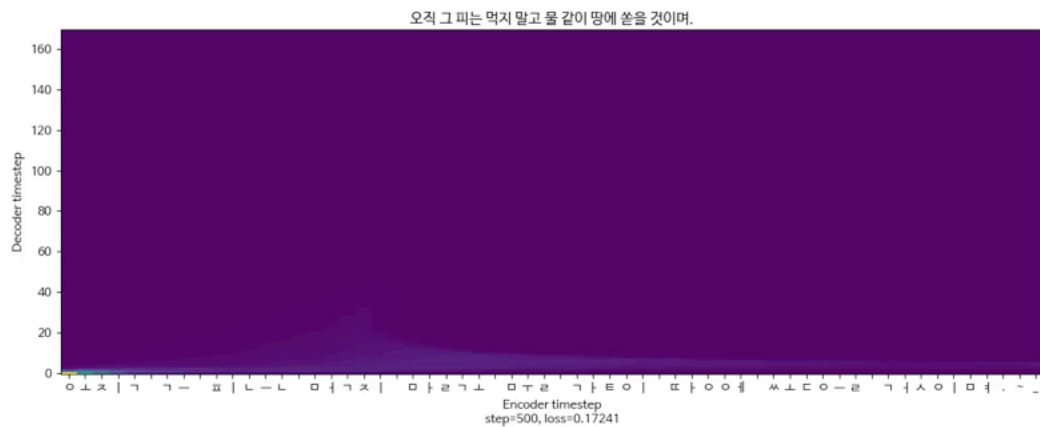
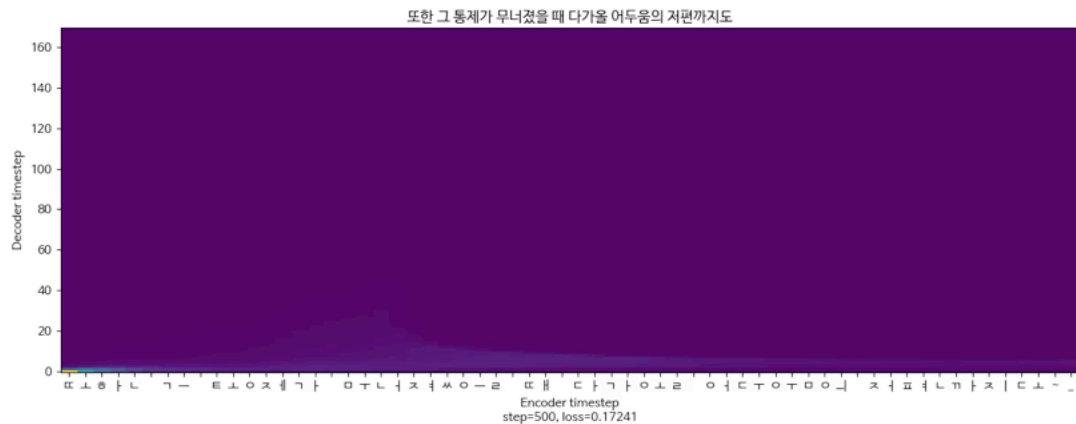
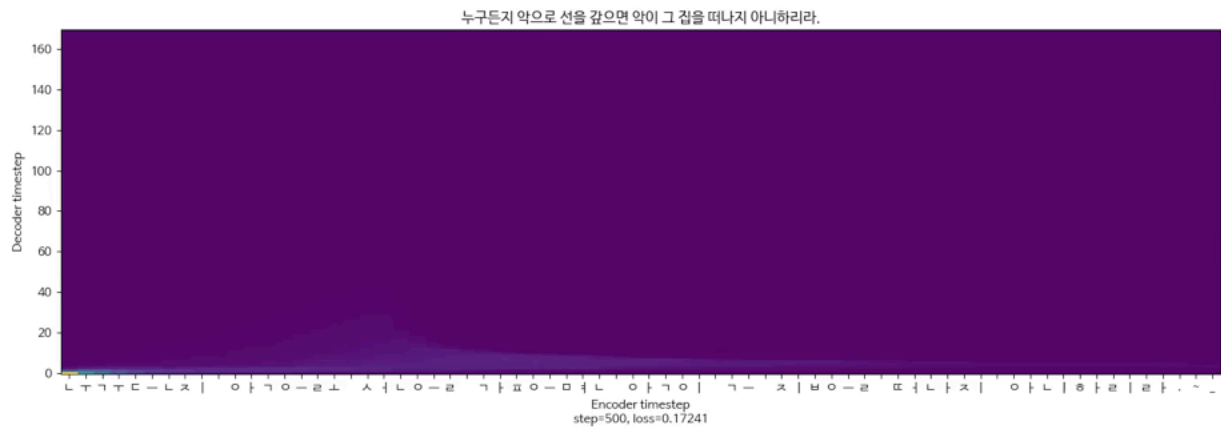
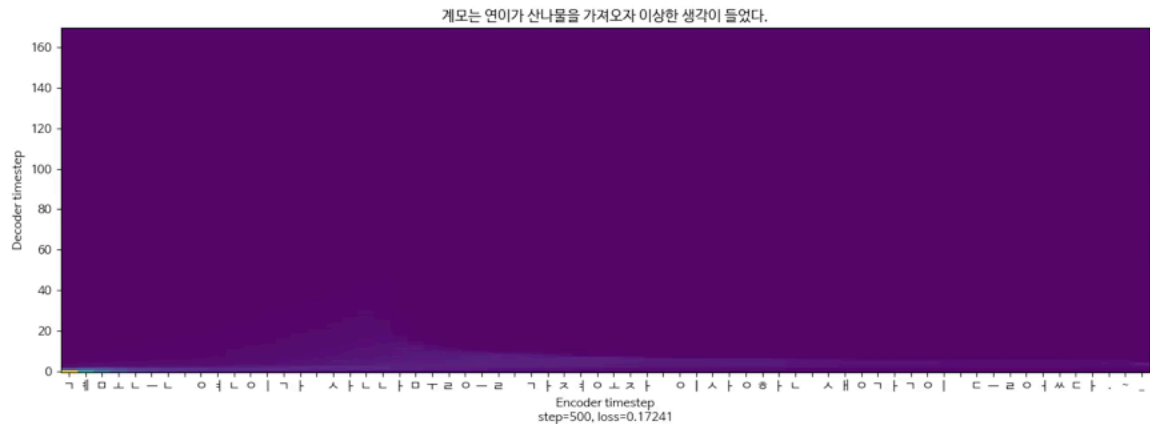
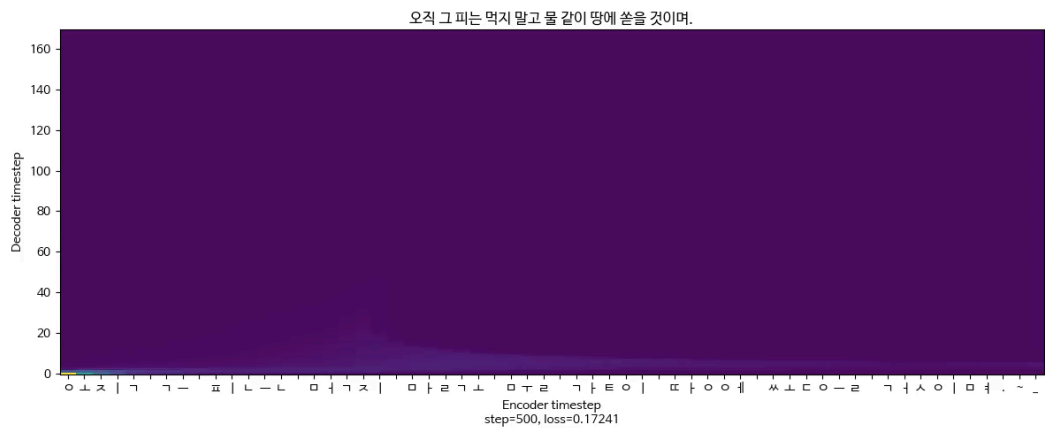
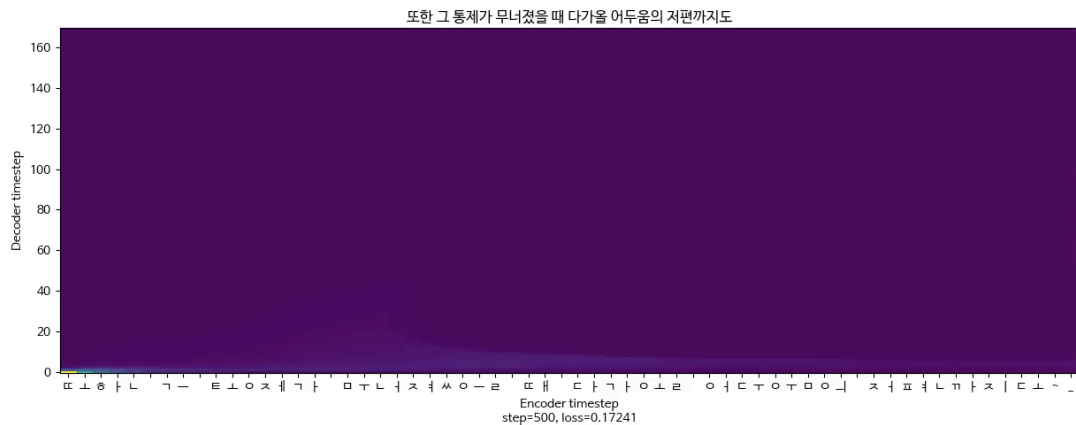
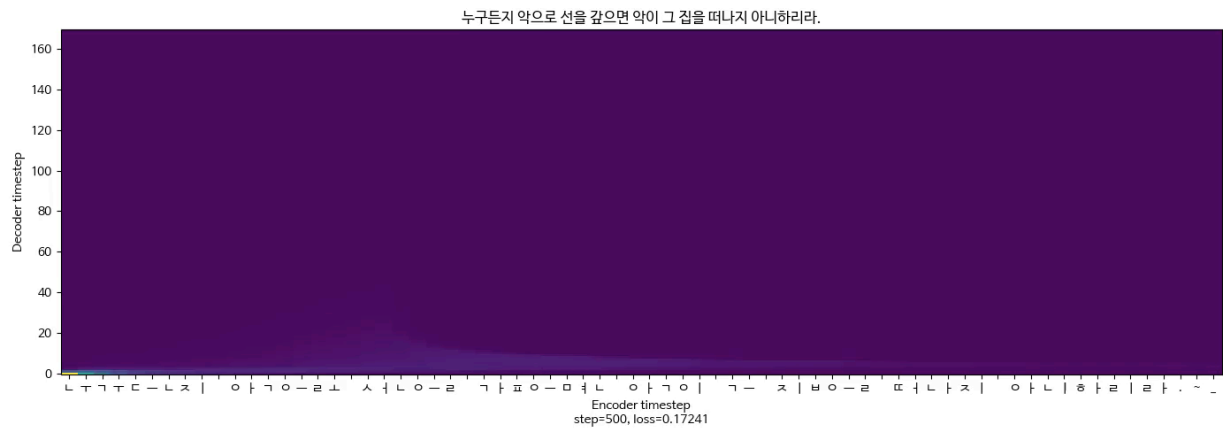
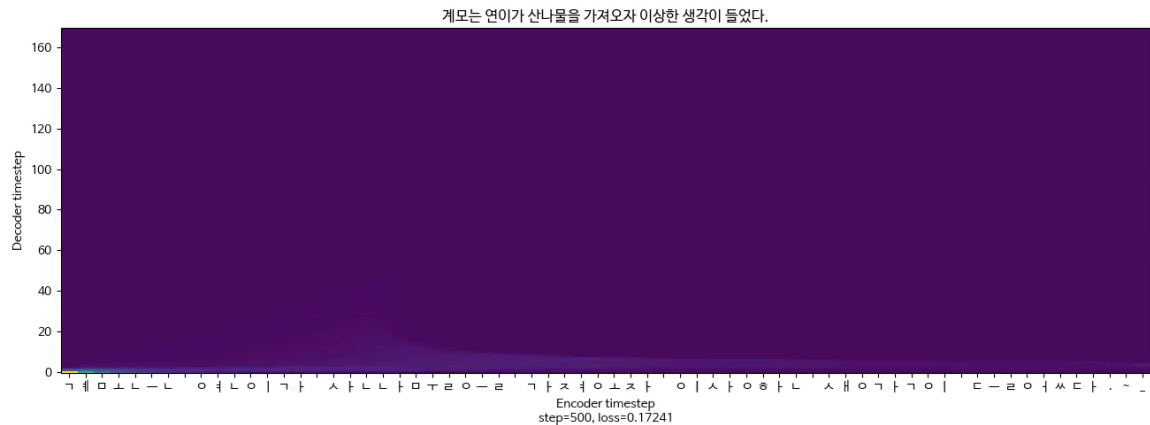
목소리 A

목소리 B

또한 그 통제가 무너졌을 때 다가올 어두움의 저편까지도

Decoder timestep

Encoder timestep

step=500, loss=0.17241

목소리 C

목소리 D

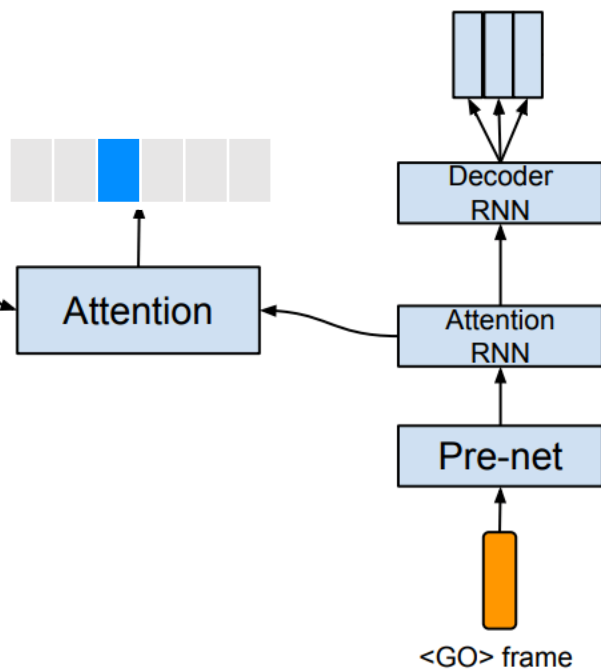
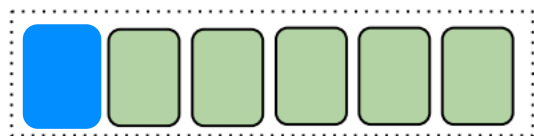
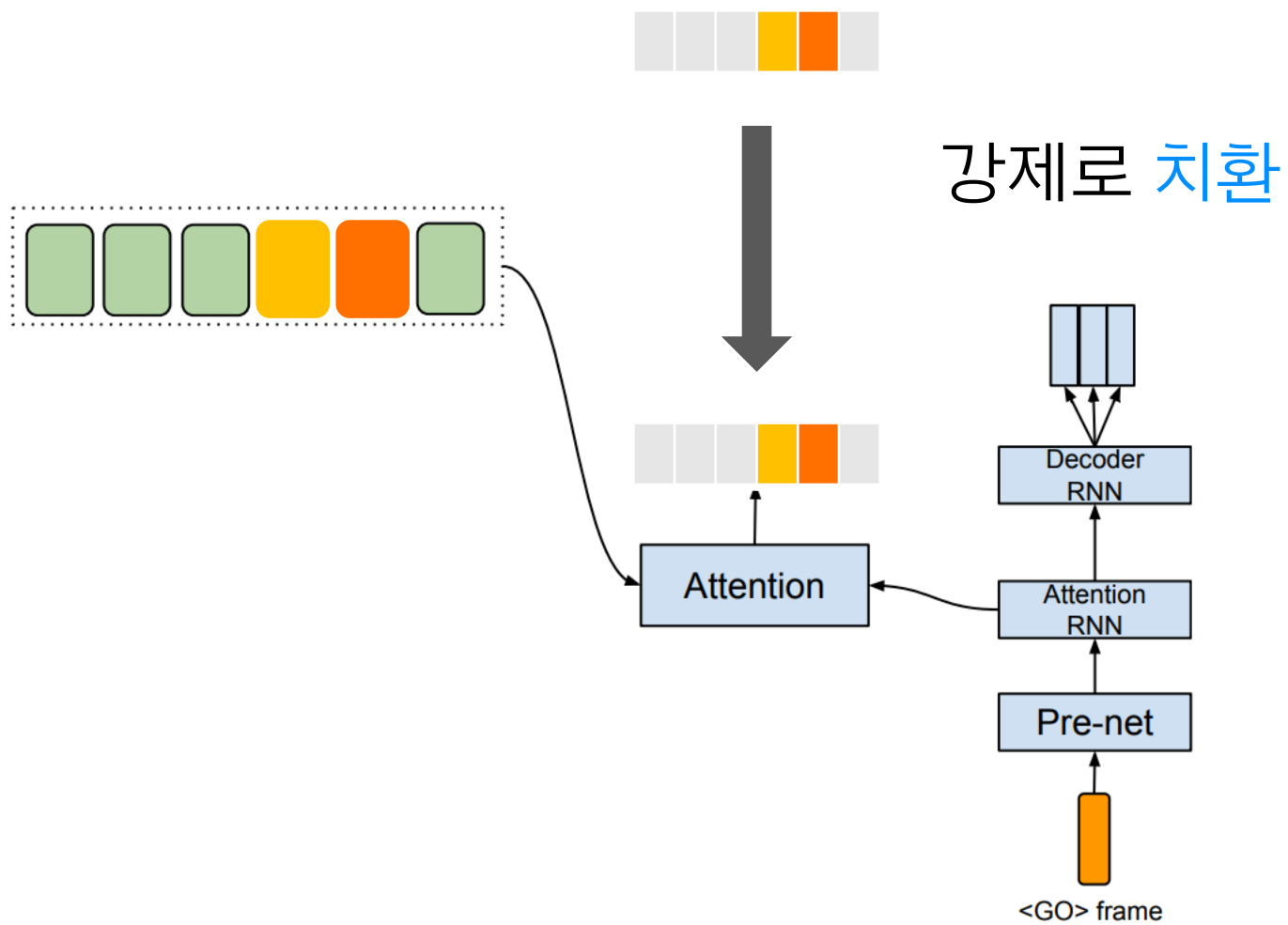
Attention을 원하는 데로 바꿔서

내가 **원하는 속도**로 말하도록 싶을 때

직접 만든 Attention

강제로 치환
Decoder RNN
Attention
Attention RNN
Pre-net
<GO> frame

Decoder timestep

Encoder timestep
[최근 딥러닝에서는 제너러티브 어드벌서리얼 네트워크가 핫하다]

한 사람의 Attention을 다른 사람에게

앵커 손석희가 문재인 대통령처럼?

성대모사?

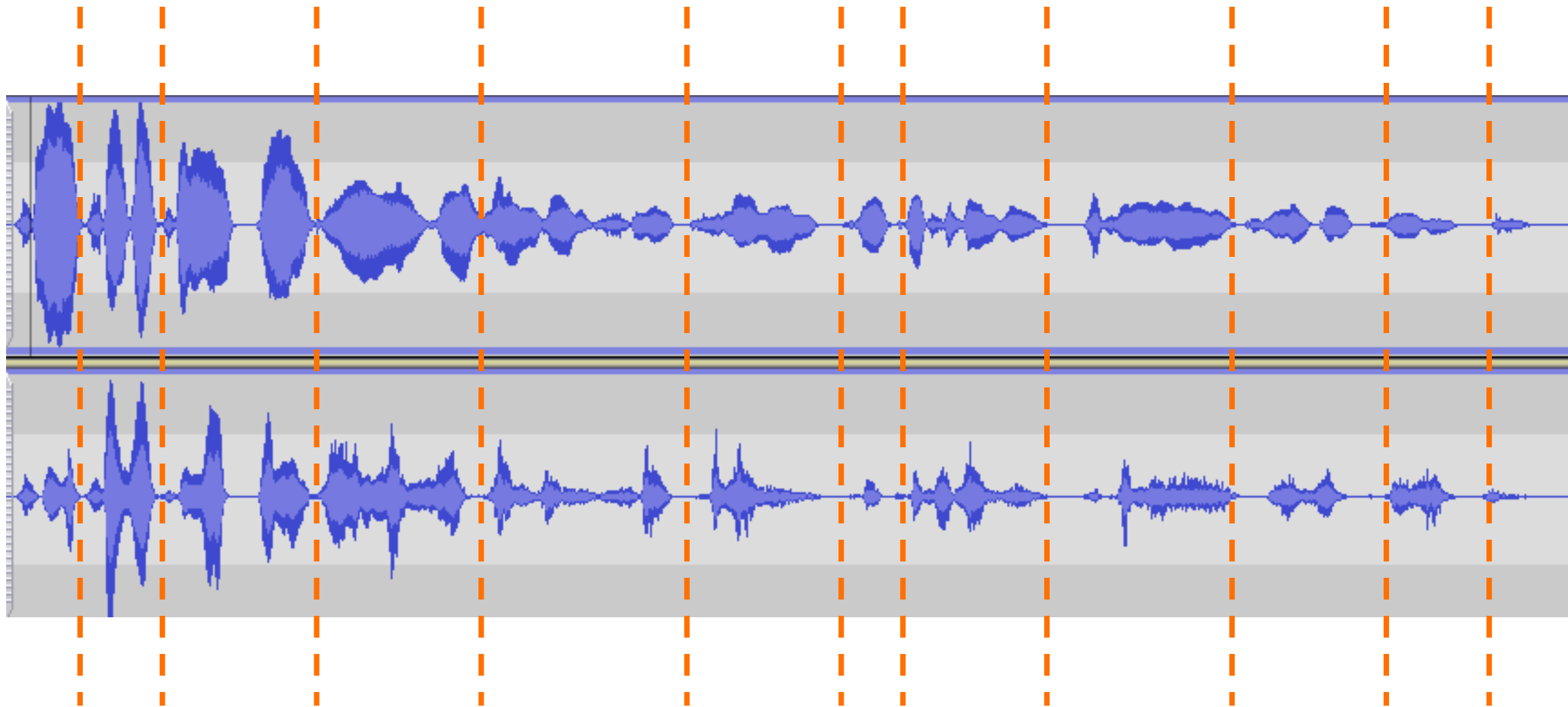
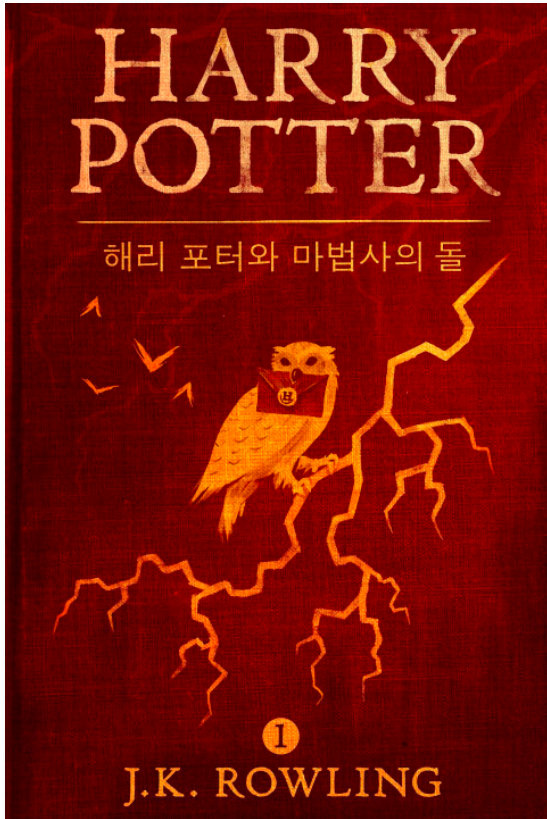
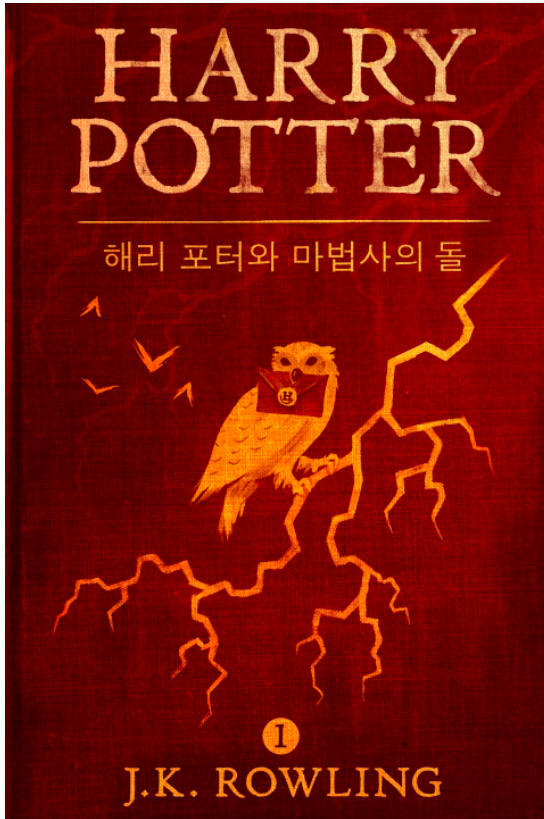
1. 개요

2. 모델

3. 결과

DEVSISTERS

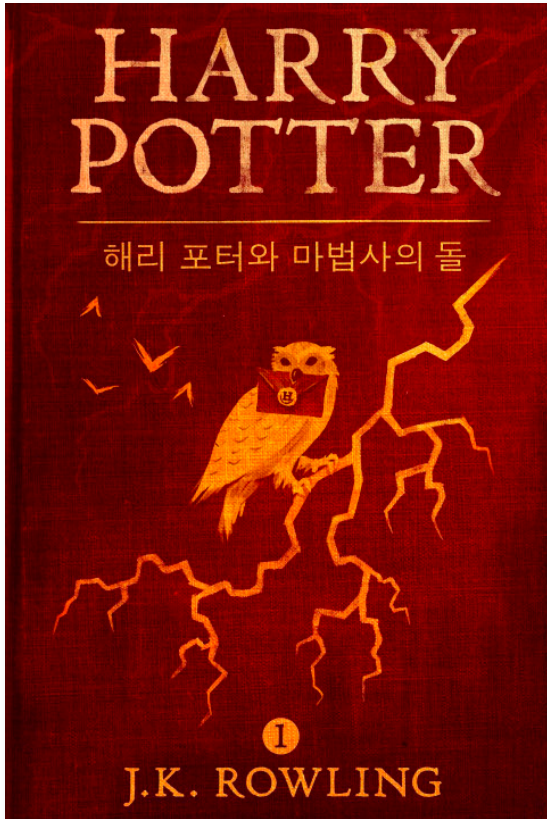
오디오북

DEVSISTERS

해리 포터와 마법사의 돌 (조앤.K.롤링 저, Pottermore from J.K. Rowling 출판) 발췌

그것은 은빛 라이터처럼 보였다.

해리 포터와 마법사의 돌 (조앤.K.롤링 저, Pottermore from J.K. Rowling 출판) 발췌

그것은 은빛 라이터처럼 보였다.

그는 뚜껑을 탁 하고 열더니 공중으로 치켜올려 찰깍 소리가 나게 했다.

해리 포터와 마법사의 돌 (조앤.K.롤링 저, Pottermore from J.K. Rowling 출판) 발췌

그것은 은빛 라이터처럼 보였다.

그는 뚜껑을 탁 하고 열더니 공중으로 치켜올려 찰칵 소리가 나게 했다.

그러자 조그맣게 펑 하는 소리가 나면서 가장 가까이 있는 가로등이 꺼졌다.

해리 포터와 마법사의 돌 (조앤.K.롤링 저, Pottermore from J.K. Rowling 출판) 발췌

박망박박투한
망박박투뭉글
토망희뿌붙날
복뷻멍멍북
볼멍벍멍터

한글 프로그래밍 언어

그리고,

kakao

연예 스포츠 KBO리그

뉴스홈 주요뉴스 이슈 랭킹뉴스 TV 뉴스톱계

9.3°C 전주 다음앱

일간 | 주간 | 월간 | 기자댓글 | 앵커브리핑

손석희 앵커의 댓글을 이제 음성으로 들어보세요

손석희의 앵커브리핑

[앵커브리핑] 판틴과 통영 성매매 여성..'그들이 만난 세상'

39 명이 공감했습니다

손석희앵커
2015.04.10 09:28

소설가 마크 트웨인이 말했습니다.
"인생에 가장 중요한 이틀이있는데, 하나는 세상에 태어난 날이고 다른 하나는 왜 이 세상에 왔는가를 깨닫는 날이다."
그런데 그 첫번째 날은 누구나 다 알지만, 두번째 날은 참 어려운 것 같습니다. 누구나 그 두번째 날을 만나기 위해 애쓰는게 삶인지도 모르겠습니다.
뉴스룸도 그런 면에서 똑같습니다. 저희들도 그 두번째의 날을 만나고 기억하기 위해 매일 매일 최선을 다하겠습니다.

일시에 꺼진 세월호 CCTV, 정전 때문?..고의차단 조사

220 명이 공감했습니다

손석희앵커
2014.08.27 16:56

세상은 세월호 특별법을 가지고 혼란스럽습니다. 세월호 참사의 본질은 무엇인가를 다시 생각하게 하는 요즈음입니다. 진상규명이든 보상이든 특별법을 가지고 이렇게 보혁갈등을 일으키는 사회가 정상적인가를 생각해봐야 하지 않을까 합니다. 세월호 참사에 대해 저희는 아직도 왜?라는 질문을 버리지 않고 있습니다. 즉, 참사의 원인에 대해 규명하지 않고서는 한 걸음도 더 나아갈 수 없다고 믿기 때문입니다. 저희는 이런 작업이 음모론에 치우치는 것도 원치 않습니다. 단지 취재한 것을 사실 그대로 전해드리는 것이 저희들의 원칙이고 그것이 저널리즘의 역할이라고 믿습니다. 오늘도 왜?라는 질문에 대해 팽목항에서 김관 기자가 전해드릴 내용들이 있습니다. 이따가 뵙겠습니다.

kakao

소설가 마크 트웨인이 말했습니다.

인생에 가장 중요한 이틀이 있는데,

하나는 세상에 태어난 날이고

다른 하나는 왜 이 세상에 왔는가를 깨닫는 날이다.

kakao

소설가 마크 트웨인이 말했습니다.

인생에 가장 중요한 이틀이 있는데,

하나는 세상에 태어난 날이고

다른 하나는 왜 이 세상에 왔는가를 깨닫는 날이다.

소설가 마크 트웨인이 말했습니다.

인생에 가장 중요한 이틀이 있는데,

하나는 세상에 태어난 날이고

다른 하나는 왜 이 세상에 왔는가를 깨닫는 날이다.

소설가 마크 트웨인이 말했습니다.

인생에 가장 중요한 이틀이 있는데,

하나는 세상에 태어난 날이고

다른 하나는 왜 이 세상에 왔는가를 깨닫는 날이다.

소설가 마크 트웨인이 말했습니다.

인생에 가장 중요한 이틀이 있는데,

하나는 세상에 태어난 날이고

다른 하나는 왜 이 세상에 왔는가를 깨닫는 날이다.

이것 외에도

아. 퇴근하고 싶다. 퇴근하고 싶다. 퇴근하고 싶다.

아. 퇴근하고 싶다. 퇴근하고 싶다. 퇴근하고 싶다.

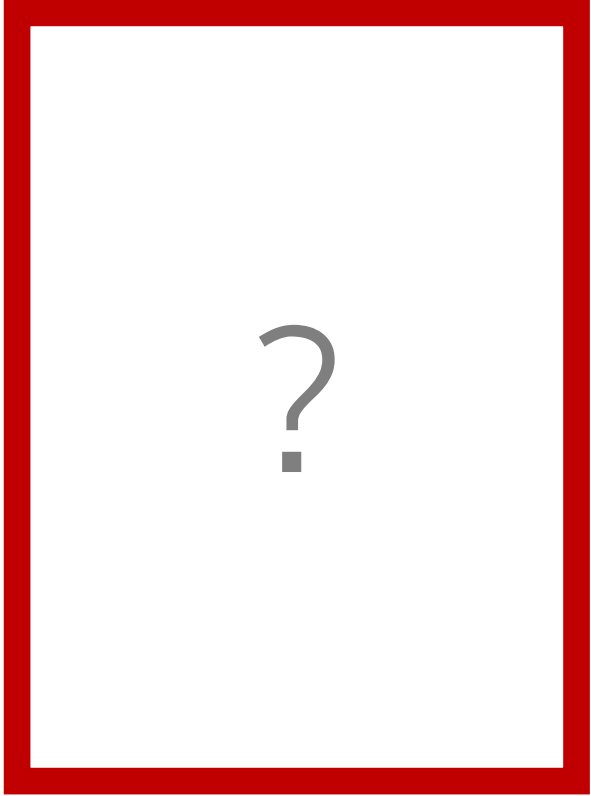
아. 퇴근하고 싶다. 퇴근하고 싶다. 퇴근하고 싶다.

시.공.조.아. × 5

아. 퇴근하고 싶다. 퇴근하고 싶다. 퇴근하고 싶다.

시.공.조.아. × 5

이겼닭! 오늘 저녁은 치킨이닭!

하지만 이런것도..

불완전한 학습 데이터 5+시간

제너러티브 어드벌서리얼 네트워크와
베리에셔널 오토 인코더가 핫하다.

기회는 평등할 것입니다.

기회는 평등할 것입니다.

과정은 공정할 것입니다.

?

기회는 평등할 것입니다.

과정은 공정할 것입니다.

결과는 정의로울 것입니다.

불완전한 학습 데이터 2+시간

네이버 웹툰 <선천적 얼간이들> - 가스파드

영웅은 죽지않아요.

영웅은 죽지않아요.

오늘 당직은 메르시입니다.

영웅은 죽지않아요.

오늘 당직은 메르시입니다.

그리고 정마담한테 주려는거 이거 이거,

이거 이거 장짜리 아니여?

네

직접 듣고 싶다면

Multi-speaker Tacotron in TensorFlow

TensorFlow implementation of **Deep Voice 2: Multi-Speaker Neural Text-to-Speech**.

Taehoon Kim
carpedm20

[GitHub (soon)] [Talk] [한국어]

Samples

(Training data = Son: 15+ hours, Park: 5+ hours, Moon: 2+ hours)

Click if you can't hear any sound

제너러티브 어드벌서리얼 네트워크와 베리에셔널 오토 인코더가 핫하다.

Seo Son Park

오스트랄로피테쿠스 아파렌시스는 멸종된 사람족 종으로, 현재에는 뼈 화석이 발견되어 있다.

Seo Son Park Moon

저는 데브시스터즈에서 머신러닝 엔지니어로 일하고 있는 김태훈입니다.

Seo Son Park Moon

http://carpedm20.github.io/tacotron

Try our engine.

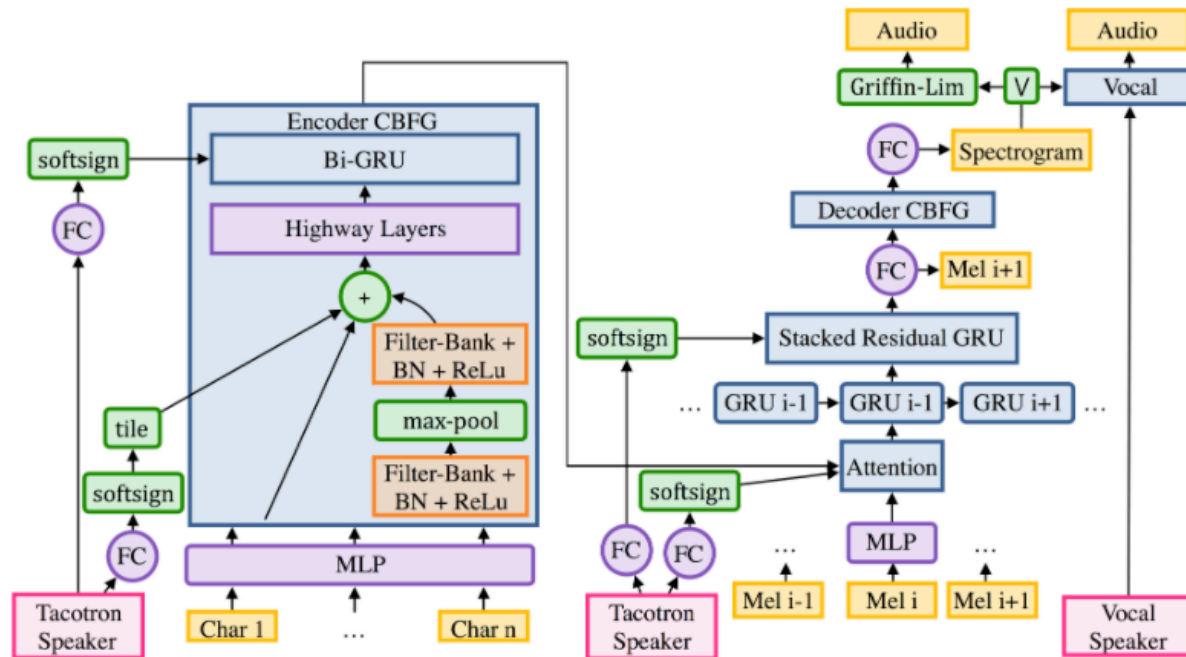
카카오뱅크가 개시 5일만에 100만 계좌를 돌파하면서 돌풍을 일으키고 있다.

Synthesize

마지막으로

마지막으로

D.Voice: 오픈소스 딥러닝 음성 합성 엔진

D.Voice는 데브시스터즈에서 개발한 오픈소스 딥러닝 음성 합성 엔진입니다.

음성 데모는 이곳에서 들어보실 수 있습니다.

사용 방법

1. 필수 라이브러리 설치

먼저 아래 명령어로 코드에 필요한 라이브러리를 설치합니다.

```
pip install -r requirements.txt
```

바로 음성을 만들고 싶으면 2-3. 학습된 모델 다운받기, 그리고 4. 음성 만들기를 따라하시면 됩니다.

```
$ python generate.py "이렇게 만드시면 됩니다"
```

코드

+

미리 학습된 모델 파라미터

+

{ 음성 추출 + 음성 자르기 + 텍스트 맞추기 }

모두 힘을 모아 유인나 데이터를...

DEVSISTERS

http://www.devsisters.com/jobs/

감사합니다

@carpedm20